

Topologia Algebrica e Percezione Visiva: come il cervello computa forme

Introduzione: L'Enigma della Costanza Percettiva

La capacità del sistema nervoso umano di riconoscere e identificare oggetti tridimensionali indipendentemente dall'angolazione di osservazione, dall'illuminazione ambientale, dalla distanza di visualizzazione e da molteplici altre trasformazioni geometriche è uno dei problemi più complessi delle neuroscienze cognitive contemporanee. Quando osserviamo una tazza da caffè sul tavolo della cucina, la riconosciamo immediatamente come tale sia che la guardiamo frontalmente, lateralmente, dall'alto o dal basso, nonostante le proiezioni retiniche risultanti da queste diverse prospettive siano radicalmente differenti dal punto di vista della geometria euclidea tradizionale. Questa straordinaria capacità, che definiamo invarianza percettiva o costanza della forma, costituisce un esempio paradigmatico di come il cervello umano estragga proprietà strutturali profonde dagli stimoli sensoriali, trascendendo le variazioni superficiali dei dati grezzi.

Il problema dell'invarianza percettiva non è semplicemente una curiosità accademica riservata agli specialisti della visione computazionale o della psicofisica sperimentale. Al contrario, esso tocca questioni fondamentali relative alla natura stessa della rappresentazione mentale, alla struttura delle informazioni sensoriali e ai principi computazionali che governano l'elaborazione neurale. Per decenni, neuroscienziati, psicologi cognitivi, matematici applicati e ingegneri hanno cercato di comprendere quali siano i meccanismi sottostanti che permettono al cervello di costruire rappresentazioni degli oggetti che siano robuste rispetto a trasformazioni come rotazioni, traslazioni, cambiamenti di scala, deformazioni elastiche e variazioni nell'illuminazione. Le teorie classiche della percezione visiva, da quelle gestaltiche a quelle computazionali sviluppate da David Marr negli anni Ottanta del secolo scorso, hanno fornito intuizioni preziose ma spesso incomplete riguardo ai fondamenti matematici di questo processo.

È in questo contesto intellettuale che la topologia algebrica emerge come un framework teorico particolarmente promettente e potente per affrontare il problema della percezione delle forme. La topologia, branca della matematica che studia le proprietà degli spazi che rimangono invarianti sotto trasformazioni continue, offre infatti un linguaggio formale naturale per descrivere caratteristiche strutturali degli oggetti che persistono attraverso deformazioni, stiramenti e piegamenti, purché non vi siano lacerazioni o incollature. A differenza della geometria euclidea classica, che si concentra su proprietà metriche come distanze, angoli e aree, la topologia si occupa di proprietà qualitative come la connessione, il numero di buchi, la compattezza e l'orientabilità. Queste proprietà topologiche risultano particolarmente rilevanti per la percezione visiva proprio perché tendono a rimanere stabili attraverso le trasformazioni che tipicamente intervengono quando osserviamo oggetti tridimensionali da diverse prospettive.

Negli ultimi due decenni, un particolare strumento della topologia algebrica applicata, noto come omologia persistente, ha rivoluzionato il modo in cui possiamo analizzare quantitativamente le strutture geometriche complesse presenti nei dati sperimentali. Sviluppata inizialmente da matematici come Herbert Edelsbrunner, Gunnar Carlsson e altri ricercatori alla fine degli anni Novanta e nei primi anni Duemila, l'omologia persistente fornisce un metodo sistematico per estrarre caratteristiche topologiche multi-scala da insiemi di punti, immagini, nuvole di dati e altri oggetti matematici complessi. Questo approccio ha trovato applicazioni in domini sorprendentemente diversi, dall'analisi delle proteine alla classificazione delle forme delle galassie, dalla previsione delle serie temporali finanziarie allo studio delle reti neurali artificiali. La sua applicazione allo studio della percezione visiva biologica rappresenta un'estensione naturale e particolarmente feconda di questa metodologia.

La Geometria della Visione: Dal Mondo Tridimensionale alla Proiezione Retinica

Per comprendere appieno il problema computazionale che il cervello deve risolvere durante il riconoscimento visivo degli oggetti, è necessario analizzare in dettaglio la geometria del processo di formazione delle immagini retiniche. Quando la luce riflessa o emessa da un oggetto tridimensionale entra nel sistema ottico dell'occhio, essa viene focalizzata sulla retina attraverso un processo che può essere approssimato, in prima istanza, come una proiezione prospettica. Questa proiezione trasforma le coordinate spaziali tridimensionali dei punti dell'oggetto in coordinate bidimensionali sulla superficie curva della retina, secondo relazioni matematiche che dipendono dalla posizione dell'oggetto rispetto all'osservatore e dalle proprietà ottiche del cristallino.

Formalmente, considerando un modello semplificato di camera a foro stenopeico (pinhole camera), la proiezione prospettica può essere descritta attraverso equazioni della forma $x_i = f * X_i / Z$ e $y_i = f * Y_i / Z$, dove (X_i, Y_i, Z_i) rappresentano le coordinate tridimensionali di un punto dell'oggetto nel sistema di riferimento della camera, (x_i, y_i) sono le coordinate bidimensionali del punto proiettato sul piano dell'immagine, f indica la distanza focale del sistema ottico e Z rappresenta la profondità del punto rispetto al piano principale della camera. Questa trasformazione geometrica è fortemente non lineare e introduce ambiguità fondamentali nel processo di percezione. In particolare, infiniti punti tridimensionali giacenti lungo uno stesso raggio visivo proiettano nello stesso punto bidimensionale dell'immagine retinica, rendendo il problema del recupero della struttura tridimensionale a partire da una singola immagine intrinsecamente sottodeterminato.

La situazione diventa ancora più complessa quando consideriamo che il cervello deve affrontare non solo questa ambiguità fondamentale ma anche una molteplicità di variazioni sistematiche nelle immagini retiniche causate da trasformazioni dell'oggetto o dell'osservatore. Una rotazione tridimensionale dell'oggetto può essere descritta matematicamente attraverso matrici di rotazione R appartenenti al gruppo speciale ortogonale $SO(3)$, che trasformano le coordinate di ogni punto secondo la relazione $X' = R * X$. Quando questa rotazione viene combinata con la proiezione prospettica, il risultato è una trasformazione altamente non lineare dell'immagine retinica che può alterare drasticamente l'aspetto bidimensionale dell'oggetto. Similmente, una traslazione dell'oggetto nello spazio

tridimensionale modifica la posizione, la scala e potenzialmente la forma apparente della proiezione retinica in modi che dipendono dalla direzione e dall'ampiezza dello spostamento.

I cambiamenti di scala rappresentano un'altra classe importante di trasformazioni che il sistema visivo deve gestire. Quando un oggetto si avvicina o si allontana dall'osservatore, la sua proiezione retinica si ingrandisce o si rimpicciolisce proporzionalmente all'inverso della distanza. Questa variazione di scala è descritta da un fattore di magnificazione $m = f / Z$, dove Z è la distanza dell'oggetto. Tuttavia, questa relazione non è semplicemente una moltiplicazione uniforme delle coordinate retiniche, poiché la proiezione prospettica introduce effetti di scorcio (foreshortening) che fanno sì che parti dell'oggetto più lontane appaiano relativamente più piccole rispetto a parti più vicine, anche all'interno dello stesso oggetto. Questo fenomeno è particolarmente evidente quando osserviamo oggetti allungati in profondità, come una strada che si allontana verso l'orizzonte o un corridoio visto frontalmente.

Oltre alle trasformazioni geometriche rigide come rotazioni, traslazioni e cambiamenti di scala, il sistema visivo deve anche confrontarsi con deformazioni non rigide degli oggetti. Gli organismi biologici, in particolare, sono caratterizzati da una notevole deformabilità: i volti cambiano espressione, i corpi assumono posture diverse, gli animali si muovono in modi complessi che alterano la configurazione spaziale delle loro parti costituenti. Anche molti oggetti inanimati possono subire deformazioni elastiche o plastiche che ne modificano la forma mantenendo però alcune proprietà strutturali fondamentali. Un esempio paradigmatico è rappresentato da un pezzo di stoffa o da un foglio di carta, che possono essere piegati, arrotolati o accartocciati in infinite configurazioni diverse pur conservando proprietà topologiche essenziali come la connessione e l'assenza di buchi.

La teoria classica della percezione visiva, particolarmente nella formulazione proposta da David Marr nel suo influente libro "Vision" pubblicato nel 1982, concettualizza il processo di riconoscimento degli oggetti come una sequenza di rappresentazioni computazionali sempre più astratte. Secondo questo modello, il sistema visivo parte da una rappresentazione primaria dell'immagine retinica, procede attraverso l'estrazione di caratteristiche locali come bordi, orientazioni e disparità binoculari per costruire quello che Marr chiama lo "schizzo primordiale" (primal sketch), quindi elabora una rappresentazione bidimensionale e mezza (two-and-a-half-dimensional sketch) che codifica informazioni sulla profondità relativa e l'orientazione delle superfici visibili, per arrivare infine a una rappresentazione tridimensionale completa dell'oggetto centrata sull'oggetto stesso piuttosto che sul punto di vista dell'osservatore.

Sebbene questo framework teorico abbia avuto un impatto enorme sul campo della visione computazionale e delle neuroscienze cognitive, esso lascia aperte numerose questioni fondamentali riguardo ai meccanismi matematici precisi attraverso cui queste rappresentazioni vengono costruite e, soprattutto, riguardo a come esse acquisiscano la proprietà dell'invarianza rispetto alle trasformazioni. In particolare, rimane poco chiaro quali siano esattamente le caratteristiche strutturali degli oggetti che vengono preservate attraverso le diverse fasi di elaborazione e che permettono il riconoscimento robusto. È qui che la topologia algebrica, e in particolare l'omologia persistente, offre nuove prospettive analitiche e computazionali.

Fondamenti di Topologia Algebrica: Spazi, Omeomorfismi e Invarianti

Prima di addentrarci nelle applicazioni specifiche della topologia algebrica alla percezione visiva, è essenziale stabilire con precisione i concetti matematici fondamentali che costituiscono il substrato teorico di questo approccio. La topologia come disciplina matematica nasce alla fine del diciannovesimo secolo e all'inizio del ventesimo secolo attraverso i contributi seminali di matematici come Henri Poincaré, L.E.J. Brouwer, Emmy Noether e molti altri. A differenza della geometria euclidea, che si occupa di proprietà metriche come lunghezze, angoli e aree, la topologia studia proprietà degli spazi che rimangono invarianti sotto trasformazioni continue e biunivoche.

Formalmente, uno spazio topologico è definito come una coppia (X, τ) dove X è un insieme e τ è una collezione di sottoinsiemi di X , chiamati aperti, che soddisfano tre assiomi fondamentali. Primo, l'insieme vuoto $\emptyset$ e l'intero insieme X devono appartenere a τ . Secondo, l'unione arbitraria di insiemi appartenenti a τ deve ancora appartenere a τ . Terzo, l'intersezione finita di insiemi appartenenti a τ deve appartenere a τ . Questa struttura astratta permette di definire rigorosamente concetti intuitivi come continuità, connessione, compattezza e convergenza senza fare riferimento a distanze metriche.

Una funzione $f: X \rightarrow Y$ tra due spazi topologici (X, τ_X) e (Y, τ_Y) è detta continua se per ogni aperto U in τ_Y , la controimmagine $f^{-1}(U)$ appartiene a τ_X . Due spazi topologici X e Y sono detti omeomorfi se esiste una funzione biunivoca $f: X \rightarrow Y$ tale che sia f che la sua inversa f^{-1} sono continue. L'omeomorfismo rappresenta la nozione fondamentale di equivalenza in topologia: spazi omeomorfi sono considerati essenzialmente identici dal punto di vista topologico, indipendentemente dalle loro differenze geometriche metriche. Questa è l'idea chiave che rende la topologia particolarmente rilevante per lo studio della percezione: due oggetti che hanno la stessa forma topologica condividono proprietà strutturali profonde che possono essere riconosciute dal sistema visivo anche quando le loro apparenze geometriche superficiali differiscono significativamente.

Un esempio classico che illustra questa distinzione è la famosa battuta secondo cui un topologo non può distinguere una tazza da una ciambella. Entrambi questi oggetti, considerati come superfici bidimensionali immerse nello spazio tridimensionale, sono omeomorfi: ciascuno ha esattamente un buco e può essere continuamente deformato nell'altro attraverso stiramenti e piegamenti senza tagli o incollature. Dal punto di vista topologico, sono lo stesso oggetto. Tuttavia, dal punto di vista geometrico metrico, presentano forme molto diverse con curvature, dimensioni e proporzioni completamente differenti. Questa invarianza topologica sotto deformazioni continue cattura esattamente il tipo di stabilità strutturale che sembra caratterizzare molti aspetti della percezione visiva umana.

Per studiare in modo sistematico le proprietà topologiche degli spazi, i matematici hanno sviluppato strumenti algebrici sofisticati che associano a ogni spazio topologico gruppi abeliani, anelli, moduli e altre strutture algebriche che codificano informazioni sulla struttura topologica dello spazio. Questi invarianti algebrici hanno la proprietà cruciale di essere preservati sotto omeomorfismi: se due spazi sono omeomorfi, i loro invarianti algebrici devono essere isomorfi. Sebbene il viceversa non sia generalmente vero, cioè spazi con invarianti isomorfi non sono necessariamente omeomorfi, questi invarianti forniscono

comunque strumenti potenti per distinguere spazi topologicamente diversi e per classificare spazi in classi di equivalenza.

Il più importante di questi invarianti algebrici è l'omologia, introdotta da Poincaré nei suoi studi pionieristici sulla topologia delle varietà. L'idea fondamentale dell'omologia è di studiare le "cavità" o "buchi" di diverse dimensioni presenti in uno spazio topologico attraverso un approccio algebrico basato sulla teoria dei gruppi. Per costruire i gruppi di omologia, si parte da una triangolazione o, più generalmente, da una struttura simpliciale sullo spazio. Un complesso simpliciale è una collezione di semplici, dove un k -simpleso è la generalizzazione in dimensione k del concetto di triangolo: un punto è un 0-simpleso, un segmento è un 1-simpleso, un triangolo è un 2-simpleso, un tetraedro è un 3-simpleso, e così via.

Formalmente, un complesso simpliciale astratto K su un insieme di vertici V è una collezione di sottoinsiemi finiti di V , chiamati semplici, tale che se σ appartiene a K e τ è un sottoinsieme di σ , allora anche τ appartiene a K . Questa condizione garantisce che il complesso contenga tutte le facce di ogni semplice. La dimensione di un semplice σ è definita come $|\sigma| - 1$, dove $|\sigma|$ indica la cardinalità dell'insieme σ . La dimensione del complesso simpliciale è il massimo delle dimensioni dei suoi semplici.

Dato un complesso simpliciale K , possiamo definire per ogni dimensione k il gruppo abeliano libero $C_k(K)$ generato dai k -semplici di K . Gli elementi di $C_k(K)$ sono combinazioni lineari formali dei k -semplici con coefficienti in un anello, tipicamente l'anello degli interi $\mathbb{Z}$ o un campo come $\mathbb{Z}_2$. Questi gruppi sono chiamati gruppi delle k -catene. La struttura fondamentale dell'omologia emerge dall'introduzione degli operatori di bordo $\partial_k: C_k(K) \rightarrow C_{k-1}(K)$, che mappano ogni k -simpleso nella somma alternata delle sue facce $(k-1)$ -dimensionali. Ad esempio, il bordo di un segmento orientato $[v_0, v_1]$ è il punto finale meno il punto iniziale: $\partial_1([v_0, v_1]) = [v_1] - [v_0]$. Il bordo di un triangolo $[v_0, v_1, v_2]$ è la somma alternata dei suoi tre lati: $\partial_2([v_0, v_1, v_2]) = [v_1, v_2] - [v_0, v_2] + [v_0, v_1]$.

La proprietà fondamentale degli operatori di bordo è che la composizione di due operatori consecutivi è sempre zero: $\partial_{k-1} \circ \partial_k = 0$ per ogni k . Questa proprietà, spesso espressa dicendo che "il bordo di un bordo è zero", non è affatto ovvia a livello intuitivo ma può essere verificata attraverso calcoli algebrici diretti. Essa implica che l'immagine di ∂_{k+1} , denotata $\text{Im}(\partial_{k+1})$, è contenuta nel nucleo di ∂_k , denotato $\text{Ker}(\partial_k)$. Gli elementi di $\text{Ker}(\partial_k)$ sono chiamati k -cicli e rappresentano k -catene che non hanno bordo. Gli elementi di $\text{Im}(\partial_{k+1})$ sono chiamati k -bordi e rappresentano k -catene che sono il bordo di qualche $(k+1)$ -catena.

Il k -esimo gruppo di omologia $H_k(K)$ è definito come il gruppo quoziente $\text{Ker}(\partial_k) / \text{Im}(\partial_{k+1})$. Intuitivamente, $H_k(K)$ misura i k -cicli che non sono bordi di $(k+1)$ -catene, cioè le k -cavità dello spazio che non possono essere "riempite" da $(k+1)$ -catene. Il rango del gruppo $H_k(K)$, chiamato k -esimo numero di Betti β_k , conta essenzialmente il numero di k -buchi indipendenti nello spazio. Il numero di Betti β_0 conta il numero di componenti connesse per cammini, β_1 conta il numero di "buchi unidimensionali" o tunnel, β_2 conta il numero di "cavità bidimensionali" o vuoti, e così via.

Per esempio, consideriamo il cerchio S^1 come spazio topologico. Una triangolazione minimale del cerchio consiste di almeno tre vertici collegati da tre segmenti a formare un ciclo. Il gruppo $H_0(S^1)$ è isomorfo a $\mathbb{Z}$, indicando che il cerchio è connesso (ha una sola

componente connessa). Il gruppo $H_1(S^1)$ è anch'esso isomorfo a $\mathbb{Z}$, riflettendo il fatto che il cerchio contiene un buco unidimensionale fondamentale che non può essere contratto a un punto. Tutti i gruppi di omologia superiori $H_k(S^1)$ per $k \geq 2$ sono banali (isomorfi al gruppo zero). In contrasto, una sfera bidimensionale S^2 ha $H_0(S^2) \cong \mathbb{Z}$ (una componente connessa), $H_1(S^2) = 0$ (nessun buco unidimensionale), $H_2(S^2) \cong \mathbb{Z}$ (una cavità bidimensionale interna), e $H_k(S^2) = 0$ per $k \geq 3$.

Un esempio più complesso è il toro bidimensionale T^2 , che può essere visualizzato come la superficie di una ciambella. Il toro ha $H_0(T^2) \cong \mathbb{Z}$ (una componente connessa), $H_1(T^2) \cong \mathbb{Z} \oplus \mathbb{Z}$ (due buchi unidimensionali indipendenti: uno che passa attraverso il buco centrale e uno che circonda il "tubo" del toro), e $H_2(T^2) \cong \mathbb{Z}$ (una cavità bidimensionale interna). Questi numeri di Betti $\beta_0 = 1$, $\beta_1 = 2$, $\beta_2 = 1$ caratterizzano completamente la topologia del toro in un senso preciso e permettono di distinguerlo da altre superfici come la sfera o il piano proiettivo.

La potenza dell'omologia risiede nel fatto che questi gruppi algebrici sono invarianti topologici completi per molte classi di spazi: se due complessi simpliciali hanno gruppi di omologia non isomorfi, allora gli spazi topologici sottostanti non possono essere omeomorfi. Inoltre, l'omologia può essere calcolata algebricamente attraverso tecniche di algebra lineare, rendendola uno strumento pratico per l'analisi computazionale di strutture topologiche complesse. Tuttavia, l'omologia classica, così come è stata descritta finora, presenta una limitazione significativa quando applicata all'analisi di dati reali: essa fornisce una descrizione globale della topologia di uno spazio ma non cattura informazioni sulla scala a cui le diverse caratteristiche topologiche emergono o scompaiono.

L'Omologia Persistente e l'Analisi Topologica Multi-Scala dei Dati

Dalla Topologia Classica alla Topologia Computazionale dei Dati

L'omologia persistente rappresenta un'estensione rivoluzionaria dell'omologia classica specificamente progettata per affrontare le sfide poste dall'analisi di dataset finiti, rumorosi e multi-scala tipici delle applicazioni scientifiche e ingegneristiche contemporanee. Mentre l'omologia tradizionale è stata sviluppata per studiare spazi topologici ideali e astratti, l'omologia persistente fornisce strumenti per estrarre informazioni topologiche significative da collezioni finite di punti, immagini digitali, nuvole di dati campionati da varietà immerse in spazi ad alta dimensione, e altre rappresentazioni discrete di strutture geometriche complesse. Questa transizione dalla topologia pura alla topologia applicata ai dati ha dato origine a un intero campo di ricerca noto come analisi topologica dei dati, o Topological Data Analysis in inglese, comunemente abbreviato come TDA.

Il problema fondamentale che l'omologia persistente affronta può essere formulato nel modo seguente. Supponiamo di avere un insieme finito di punti $X = \{x_1, x_2, \dots, x_n\}$ campionati da una varietà differenziabile M immersa in uno spazio euclideo $\mathbb{R}^d$. Questi punti potrebbero rappresentare, ad esempio, misurazioni sperimentali, pixel di un'immagine, coordinate di atomi in una molecola proteica, o pattern di attivazione neurale registrati da elettrodi. Il nostro obiettivo è recuperare informazioni sulla topologia della varietà sottostante

M a partire dal campione finito X . Tuttavia, questo recupero non è immediato perché un insieme finito di punti isolati ha una topologia banale: è semplicemente un'unione disgiunta di punti, con nessuna struttura interessante di connessione o cavità.

Per superare questo problema, dobbiamo in qualche modo "ispessire" i punti del dataset costruendo un complesso simpliciale o uno spazio topologico più ricco che approssimi la geometria della varietà sottostante. L'approccio classico consiste nel considerare l'unione di sfere di raggio ε centrate su ciascun punto del dataset. Formalmente, definiamo l'unione di sfere $U_\varepsilon(X) = \bigcup_{i=1}^n B(x_i, \varepsilon)$, dove $B(x_i, \varepsilon)$ denota la sfera chiusa di raggio ε centrata in x_i . Al variare del parametro di scala ε , otteniamo una famiglia di spazi topologici annidati: per $\varepsilon = 0$ abbiamo semplicemente i punti isolati, mentre per ε sufficientemente grande tutti i punti sono contenuti in un'unica sfera gigante.

Un problema computazionale con l'unione di sfere è che essa è un oggetto geometrico continuo difficile da manipolare algebricamente. Per rendere il problema trattabile, i matematici hanno introdotto complessi simpliciali discreti che approssimano l'unione di sfere in modi diversi. Il più importante di questi è il complesso di Čech, denotato $\check{C}_\varepsilon(X)$, definito come il complesso simpliciale i cui k -simplessi sono tutti i sottoinsiemi di $(k+1)$ punti la cui intersezione di sfere di raggio ε è non vuota. Formalmente, un insieme $\{x_{i_0}, x_{i_1}, \dots, x_{i_k}\}$ forma un k -simpleso in $\check{C}_\varepsilon(X)$ se e solo se l'intersezione $\bigcap_{j=0}^k B(x_{i_j}, \varepsilon) \neq \emptyset$.

Il complesso di Čech ha la proprietà teorica notevole di essere omotopicamente equivalente all'unione di sfere $U_\varepsilon(X)$, il che significa che i due spazi hanno gli stessi gruppi di omologia. Questo risultato, conseguenza del teorema dei nervi di Pavel Alexandrov, garantisce che calcolando l'omologia del complesso di Čech otteniamo esattamente l'informazione topologica contenuta nell'unione di sfere. Tuttavia, il complesso di Čech ha uno svantaggio computazionale significativo: determinare se $k+1$ punti formano un k -simpleso richiede di verificare se $k+1$ sfere hanno un'intersezione non vuota, un problema che diventa rapidamente oneroso per grandi valori di k e grandi dataset.

Per questa ragione, in pratica si utilizza spesso il complesso di Vietoris-Rips, denotato $VR_\varepsilon(X)$, che è computazionalmente più efficiente pur fornendo informazioni topologiche simili. Il complesso di Vietoris-Rips include un k -simpleso su punti $\{x_{i_0}, x_{i_1}, \dots, x_{i_k}\}$ se e solo se la distanza tra ogni coppia di questi punti è al massimo 2ε . In altre parole, $VR_\varepsilon(X)$ è il completamento simpliciale del grafo i cui vertici sono i punti di X e i cui spigoli connettono coppie di punti a distanza al massimo 2ε . Questa definizione è molto più semplice da implementare algebricamente perché richiede solo il calcolo di distanze a coppie tra punti.

Sebbene il complesso di Vietoris-Rips non sia in generale omotopicamente equivalente all'unione di sfere, esiste una relazione precisa tra i due: si può dimostrare che $\check{C}_\varepsilon(X)$ è un sottocomplesso di $VR_\varepsilon(X)$, che a sua volta è un sottocomplesso di $\check{C}_{2\varepsilon}(X)$. Questa catena di inclusioni implica che le caratteristiche topologiche presenti nel complesso di Čech a scala ε sono certamente presenti anche nel complesso di Vietoris-Rips alla stessa scala, mentre le caratteristiche presenti in quest'ultimo sono certamente presenti nel complesso di Čech a scala 2ε . In pratica, per molti dataset, i complessi di Čech e Vietoris-Rips producono risultati qualitativamente simili, rendendo il secondo una scelta prediletta per applicazioni computazionali.

Un altro complesso simpliciale ampiamente utilizzato nell'analisi topologica dei dati è il complesso alfa, che può essere pensato come una versione più raffinata del complesso di Čech che tiene conto della triangolazione di Delaunay dei punti. Il complesso alfa $\alpha_\epsilon(X)$ è definito come il sottocomplesso del complesso di Delaunay formato dai semplici il cui circonraggio (il raggio della più piccola sfera che circoscrive il semplice) è al massimo ϵ . Per insiemi di punti in spazi euclidei di bassa dimensione (tipicamente $\mathbb{R}^2$ o $\mathbb{R}^3$), il complesso alfa può essere calcolato efficientemente sfruttando algoritmi geometrici computazionali per la triangolazione di Delaunay, ed esso fornisce un'approssimazione particolarmente fedele della forma dei dati.

Indipendentemente dal tipo specifico di complesso simpliciale utilizzato, l'idea chiave dell'omologia persistente è di studiare non solo la topologia a una singola scala ϵ fissata, ma piuttosto l'evoluzione della topologia attraverso un intero intervallo di scale. Formalmente, consideriamo una filtrazione, che è una sequenza annidata di complessi simpliciali $K_0 \subseteq K_1 \subseteq K_2 \subseteq \dots \subseteq K_m$, dove ciascun complesso K_i è un sottocomplesso del successivo K_{i+1} . Nel contesto dell'analisi di dati, questa filtrazione è tipicamente costruita prendendo $K_i = VR_{\{\epsilon_i\}}(X)$ per una sequenza crescente di parametri di scala $0 = \epsilon_0 < \epsilon_1 < \epsilon_2 < \dots < \epsilon_m$.

Ogni complesso K_i nella filtrazione ha i propri gruppi di omologia $H_k(K_i)$ che descrivono le caratteristiche topologiche presenti a quella scala. Le inclusioni $K_i \subseteq K_{i+1}$ inducono omomorfismi tra i gruppi di omologia, $f_{i,i+1}: H_k(K_i) \rightarrow H_k(K_{i+1})$, che mappano le caratteristiche topologiche a scala ϵ_i nelle corrispondenti caratteristiche a scala ϵ_{i+1} . L'omologia persistente studia la struttura di questi omomorfismi per identificare quali caratteristiche topologiche persistono attraverso un ampio intervallo di scale e quali invece appaiono e scompaiono rapidamente.

Più precisamente, data una filtrazione di complessi simpliciali, possiamo tracciare il "ciclo di vita" di ogni caratteristica topologica, identificando la scala di nascita (birth) alla quale la caratteristica appare per la prima volta e la scala di morte (death) alla quale essa scompare fondendosi con un'altra caratteristica o venendo riempita. Una caratteristica che nasce a scala b e muore a scala d ha una persistenza pari a $d - b$. L'intuizione fondamentale dell'omologia persistente è che le caratteristiche con alta persistenza corrispondono a proprietà topologiche robuste e significative della struttura sottostante i dati, mentre le caratteristiche con bassa persistenza sono spesso dovute a rumore, variazioni casuali o artefatti del campionamento.

Il teorema fondamentale dell'omologia persistente, dimostrato da Zomorodian e Carlsson nel 2005, stabilisce che l'omologia persistente può essere completamente decomposta in moduli intervallo indecomponibili. In altre parole, la struttura algebrica della filtrazione può essere descritta come una somma diretta di moduli della forma $k[b, d)$, dove b indica la scala di nascita e d la scala di morte della caratteristica corrispondente. Questa decomposizione è unica a meno di permutazioni, il che significa che l'insieme di tutti gli intervalli di persistenza $[b, d)$ costituisce un invariante topologico completo della filtrazione, noto come codice a barre di persistenza (persistence barcode).

Visivamente, il codice a barre di persistenza può essere rappresentato come un diagramma in cui ogni caratteristica topologica corrisponde a un intervallo orizzontale che inizia alla

scala di nascita e termina alla scala di morte. Le caratteristiche con barre lunghe rappresentano proprietà topologiche robuste che persistono attraverso molte scale, mentre le caratteristiche con barre corte rappresentano fluttuazioni transitorie. Il codice a barre fornisce quindi un riassunto compatto ma completo dell'informazione topologica multi-scala contenuta nei dati.

Una rappresentazione alternativa ma equivalente dell'omologia persistente è il diagramma di persistenza (persistence diagram), che è un diagramma di dispersione nel piano (b, d) dove ogni caratteristica topologica è rappresentata da un punto con coordinate (nascita, morte). I punti lontani dalla diagonale $b = d$ corrispondono a caratteristiche con alta persistenza, mentre i punti vicini alla diagonale corrispondono a caratteristiche con bassa persistenza che nascono e muoiono rapidamente. La diagonale stessa, che teoricamente contiene infiniti punti corrispondenti a caratteristiche di persistenza zero, viene spesso inclusa nel diagramma con molteplicità infinita per ragioni matematiche legate alla teoria della stabilità.

Un risultato teorico cruciale che rende l'omologia persistente particolarmente adatta all'analisi di dati reali è il teorema di stabilità dimostrato da David Cohen-Steiner, Herbert Edelsbrunner e John Harer nel 2007. Questo teorema stabilisce che l'omologia persistente è stabile rispetto a piccole perturbazioni dei dati di input. Formalmente, dati due dataset X e Y che sono vicini nel senso della distanza di Hausdorff, i loro diagrammi di persistenza sono vicini nel senso della distanza di bottleneck, che misura la distanza massima che un punto di un diagramma deve percorrere per essere accoppiato con un punto dell'altro diagramma. Questa proprietà di stabilità garantisce che le caratteristiche topologiche identificate dall'omologia persistente non sono artefatti fragili del campionamento esatto ma riflettono proprietà robuste della struttura geometrica sottostante.

Algoritmi Computazionali per l'Omologia Persistente

Il calcolo effettivo dell'omologia persistente richiede algoritmi sofisticati che siano in grado di gestire complessi simpliciali potenzialmente molto grandi derivanti da dataset di migliaia o milioni di punti. Il problema computazionale fondamentale consiste nel ridurre la matrice di bordo del complesso simpliciale filtrato a una forma canonica dalla quale si possono leggere direttamente gli intervalli di persistenza. L'algoritmo standard per questa riduzione è una variante dell'eliminazione gaussiana adattata alla struttura specifica delle filtrazioni simpliciali.

L'algoritmo procede costruendo prima il complesso simpliciale completo per la scala massima di interesse, quindi ordinando tutti i semplici secondo l'ordine di apparizione nella filtrazione. Questo ordinamento deve rispettare la condizione che se un semplice σ è una faccia di un semplice τ , allora σ deve precedere τ nell'ordinamento. Un tale ordinamento è sempre possibile e non è unico, ma diverse scelte producono lo stesso risultato finale grazie all'invarianza della decomposizione in moduli intervallo.

Data una filtrazione ordinata di semplici $\sigma_1, \sigma_2, \dots, \sigma_N$, si costruisce la matrice di bordo D dove l'entrata $D[i,j]$ è il coefficiente di σ_i nella rappresentazione del bordo di σ_j . Questa matrice ha la proprietà che $D^2 = 0$, riflettendo il fatto algebrico che il bordo di un bordo è sempre zero. L'algoritmo di riduzione opera su questa matrice attraverso operazioni sulle colonne (che corrispondono ad addizioni di catene) fino a portarla in una forma ridotta dove

ogni colonna contiene al massimo un elemento non zero più in basso di tutte le altre colonne precedenti.

Nella matrice ridotta, le colonne possono essere classificate in due tipi. Le colonne positive sono quelle che contengono almeno un elemento non zero dopo la riduzione. L'elemento non zero più basso in una colonna positiva identifica la scala di morte di una caratteristica topologica, mentre la riga corrispondente identifica la scala di nascita. Le colonne negative sono quelle che risultano identicamente nulle dopo la riduzione e corrispondono alla nascita di nuove caratteristiche topologiche che non muoiono mai (o muoiono solo all'infinito se si considera la filtrazione estesa).

La complessità computazionale dell'algoritmo di riduzione dipende criticamente dal numero totale di semplici nella filtrazione, che può crescere esponenzialmente con il numero di punti nel dataset. Per un complesso di Vietoris-Rips su n punti, il numero di k -semplici può essere dell'ordine di n^{k+1} , il che rende proibitivo il calcolo dell'omologia in dimensioni elevate per grandi dataset. Questa esplosione combinatoria ha motivato lo sviluppo di numerose ottimizzazioni algoritmiche e approssimazioni.

Un'importante ottimizzazione è il calcolo dell'omologia persistente ridotta, che sfrutta il fatto che spesso siamo interessati solo alle caratteristiche topologiche in dimensioni basse (tipicamente $k = 0, 1, 2$) e possiamo quindi evitare di costruire esplicitamente i semplici di dimensione superiore. Tecniche di compressione come i witness complexes, introdotti da Vin de Silva e Gunnar Carlsson, riducono drasticamente la dimensione del problema costruendo complessi simpliciali su un sottoinsieme di punti landmark scelti strategicamente dal dataset completo.

Un'altra direzione di ricerca ha portato allo sviluppo di algoritmi di approssimazione che calcolano un'omologia persistente approssimata in tempo polinomiale rispetto al numero di punti, sacrificando una piccola quantità di precisione nella localizzazione esatta degli intervalli di persistenza. Questi algoritmi sono particolarmente preziosi per applicazioni su larga scala dove l'accuratezza perfetta non è essenziale ma la velocità computazionale è critica.

Negli ultimi anni, l'implementazione pratica dell'omologia persistente ha beneficiato enormemente dello sviluppo di librerie software efficienti e ben documentate. Il pacchetto GUDHI (Geometry Understanding in Higher Dimensions), sviluppato dall'INRIA in Francia, fornisce implementazioni ottimizzate di algoritmi per l'omologia persistente in C++ con interfacce Python. La libreria Ripser, sviluppata da Ulrich Bauer, è specificamente ottimizzata per il calcolo veloce dell'omologia persistente su complessi di Vietoris-Rips e detiene attualmente numerosi record di velocità. Altri pacchetti popolari includono Dionysus, JavaPlex e Perseus, ciascuno con i propri punti di forza per diverse applicazioni.

Vettorizzazione e Apprendimento Automatico con l'Omologia Persistente

Sebbene i codici a barre di persistenza e i diagrammi di persistenza forniscano rappresentazioni ricche e complete dell'informazione topologica multi-scala, essi presentano una sfida quando vogliamo utilizzare l'omologia persistente in congiunzione con tecniche di apprendimento automatico statistico. Gli algoritmi classici di machine learning, come le reti neurali, le macchine a vettori di supporto o gli alberi di decisione, richiedono tipicamente

input vettoriali in uno spazio euclideo di dimensione fissa. I diagrammi di persistenza, invece, sono multi-insiemi di punti in R^2 con cardinalità variabile e non hanno una rappresentazione vettoriale canonica.

Questo problema ha motivato lo sviluppo di numerosi metodi di vettorizzazione, o featurization, che trasformano i diagrammi di persistenza in vettori di caratteristiche di dimensione fissa che possono essere utilizzati come input per algoritmi di machine learning. Ciascuno di questi metodi cerca di catturare aspetti diversi dell'informazione contenuta nel diagramma di persistenza, spesso sacrificando la completezza dell'informazione in cambio di maggiore trattabilità computazionale e compatibilità con framework di apprendimento esistenti.

Una delle prime e più semplici tecniche di vettorizzazione è il vettore delle statistiche sommarie (summary statistics vector), che calcola statistiche descrittive sui punti del diagramma di persistenza. Esempi tipici includono il numero totale di punti in ciascuna dimensione omologica, la persistenza media, la persistenza massima, la varianza della persistenza, il tempo di nascita medio, e così via. Sebbene questo approccio sia computazionalmente efficiente e fornisca vettori di dimensione bassa facili da interpretare, esso ovviamente scarta molta dell'informazione dettagliata contenuta nella distribuzione spaziale completa dei punti nel diagramma.

Un approccio più sofisticato è il paesaggio di persistenza (persistence landscape), introdotto da Peter Bubenik nel 2015. Il paesaggio di persistenza è una famiglia di funzioni continue $\lambda_k: R \rightarrow R$, per $k = 1, 2, 3, \dots$, costruite a partire dal diagramma di persistenza in modo da catturare la struttura completa della persistenza. Formalmente, per ogni punto t sulla scala dei parametri, si considerano tutti gli intervalli di persistenza $[b_i, d_i]$ e si calcola la funzione tenda centrata su ciascun intervallo: $f_i(t) = \min(t - b_i, d_i - t, 0)$. Il k -esimo paesaggio di persistenza $\lambda_k(t)$ è definito come il k -esimo valore più alto tra tutte le funzioni tenda valutate in t .

I paesaggi di persistenza hanno diverse proprietà matematiche attraenti. Primo, essi formano uno spazio vettoriale, il che significa che possono essere sommati e moltiplicati per scalari in modo naturale. Secondo, la mappa dal diagramma di persistenza al paesaggio di persistenza è iniettiva, cioè diagrammi diversi producono paesaggi diversi, garantendo che nessuna informazione essenziale viene persa nella trasformazione. Terzo, i paesaggi godono di una proprietà di stabilità analoga a quella dei diagrammi di persistenza: piccole perturbazioni nel diagramma producono piccole perturbazioni nel paesaggio secondo una norma funzionale appropriata. Quarto, i paesaggi possono essere discretizzati campionandoli su una griglia uniforme, producendo vettori di dimensione fissa che possono essere utilizzati direttamente negli algoritmi di machine learning.

Un altro metodo di vettorizzazione ampiamente utilizzato è l'immagine di persistenza (persistence image), introdotta da Henry Adams e colleghi nel 2017. L'idea è di pensare al diagramma di persistenza come a una distribuzione di massa nel piano (nascita, morte) e di rappresentare questa distribuzione attraverso un'immagine discreta bidimensionale. Il metodo opera in due fasi. Prima, ogni punto (b, d) del diagramma viene trasformato in una funzione gaussiana bidimensionale centrata in quel punto, con varianza controllata da un parametro di smoothing. Secondo, lo spazio viene diviso in una griglia di pixel e l'intensità di

ciascun pixel è calcolata integrando la somma di tutte le gaussiane su quel pixel. Il risultato è un'immagine discretizzata che può essere vettorizzata semplicemente concatenando i valori dei pixel in un unico vettore.

Le immagini di persistenza hanno il vantaggio di essere estremamente intuitive e facili da visualizzare, e possono essere calcolate in modo computazionalmente efficiente. Inoltre, esse sono naturalmente compatibili con architetture di deep learning progettate per processare immagini, come le reti neurali convoluzionali. Tuttavia, a differenza dei paesaggi di persistenza, la trasformazione in immagini di persistenza non è iniettiva, il che significa che diagrammi leggermente diversi possono produrre la stessa immagine se i loro punti cadono nello stesso pixel dopo la discretizzazione.

Neurobiologia della Visione e Rappresentazioni Invarianti

Architettura del Sistema Visivo dei Mammiferi

Per comprendere come l'omologia persistente e gli strumenti della topologia algebrica possano illuminare i meccanismi della percezione visiva, è necessario prima esaminare in dettaglio l'organizzazione anatomica e funzionale del sistema visivo dei mammiferi, con particolare attenzione agli esseri umani e ai primati non umani che sono stati oggetto di studi neuroscientifici estensivi. Il percorso dell'informazione visiva dal mondo esterno alla rappresentazione corticale è un processo straordinariamente complesso che coinvolge trasformazioni successive attraverso molteplici stadi neurali, ciascuno dei quali estrae e trasforma l'informazione in modi specifici.

Il processo inizia nella retina, un tessuto neurale stratificato di circa mezzo millimetro di spessore che riveste la superficie interna posteriore del globo oculare. La retina contiene circa centotrenta milioni di fotorecettori, suddivisi in bastoncelli, responsabili della visione in condizioni di scarsa illuminazione, e coni, responsabili della visione dei colori e della visione ad alta acuità in condizioni di buona illuminazione. Quando i fotoni di luce colpiscono le molecole di fotopigmento contenute nei segmenti esterni dei fotorecettori, innescano una cascata di reazioni biochimiche che portano all'iperpolarizzazione della membrana cellulare e alla modulazione del rilascio di neurotrasmettitori nelle sinapsi con le cellule bipolari, il secondo strato neurale della retina.

Le cellule bipolari ricevono input da multipli fotorecettori attraverso un pattern di connettività che crea campi recettivi centro-periferia. Un campo recettivo è definito come la regione dello spazio visivo all'interno della quale stimoli luminosi possono modulare l'attività di un neurone. Nel caso delle cellule bipolari, esistono due principali tipi di organizzazione: cellule ON che aumentano la loro attività quando il centro del loro campo recettivo è illuminato e la periferia è al buio, e cellule OFF che mostrano il comportamento opposto. Questa organizzazione antagonista centro-periferia, scoperta da Stephen Kuffler negli anni Cinquanta, rappresenta il primo stadio di elaborazione dell'informazione visiva e implementa essenzialmente un'operazione di rilevamento dei contrasti locali nell'immagine retinica.

Le cellule bipolari proiettano alle cellule gangliari retiniche, il terzo e ultimo strato neurale della retina, i cui assoni formano il nervo ottico che trasporta i segnali visivi dalla retina al cervello. Esistono diversi tipi funzionali di cellule gangliari, ma i due più abbondanti sono le cellule gangliari M, o magnocellulari, che rispondono preferenzialmente a stimoli in movimento e a contrasti di luminanza e hanno risoluzioni temporali elevate, e le cellule gangliari P, o parvocellulari, che hanno campi recettivi più piccoli, risoluzioni spaziali più alte e contribuiscono maggiormente alla visione dei colori. Oltre a questi tipi principali, esistono numerosi altri sottotipi di cellule gangliari specializzate per rilevare caratteristiche specifiche come la direzione del movimento, il colore locale, o variazioni nell'illuminazione globale.

Gli assoni delle cellule gangliari retiniche formano il nervo ottico che si proietta principalmente al nucleo genicolato laterale del talamo, una struttura sottocorticale che funge da stazione di rilascio per la maggior parte dell'informazione visiva diretta alla corteccia. Il nucleo genicolato laterale è organizzato in sei strati principali nei primati: gli strati uno e due, chiamati magnocellulari, ricevono input dalle cellule gangliari M e processano informazioni relative al movimento e ai contrasti di luminanza; gli strati da tre a sei, chiamati parvocellulari, ricevono input dalle cellule gangliari P e sono coinvolti nella visione ad alta risoluzione spaziale e nella discriminazione dei colori. Esiste anche una serie di strati intercalati tra gli strati principali, chiamati strati koniocellulari, che processano informazioni cromatiche specifiche provenienti da sottotipi specializzati di cellule gangliari.

Sorprendentemente, sebbene il nucleo genicolato laterale riceva input massiccio dalla retina e proietti massicciamente alla corteccia visiva primaria, la maggior parte delle sinapsi che le cellule genicolate ricevono proviene non dalla retina ma da feedback corticali e da altre regioni cerebrali. Questo suggerisce che il nucleo genicolato laterale non è semplicemente una stazione di passaggio passiva ma svolge un ruolo attivo nel modulare e filtrare l'informazione visiva in base al contesto comportamentale, all'attenzione e ad altri fattori top-down. Tuttavia, dal punto di vista della trasformazione dell'informazione sensoriale, i neuroni genicolati mantengono campi recettivi centro-periferia simili a quelli delle cellule gangliari che li innervano, suggerendo che non avviene un'elaborazione radicalmente nuova a questo stadio.

L'informazione visiva dal nucleo genicolato laterale viene trasmessa alla corteccia visiva primaria, anche chiamata V1 o area 17 secondo la nomenclatura di Brodmann, situata nel lobo occipitale nella parte posteriore del cervello. La corteccia V1 è la prima area corticale a ricevere input visivi e rappresenta il sito dove avviene una trasformazione profonda della rappresentazione dell'informazione. Mentre le cellule della retina e del genicolato hanno campi recettivi circolari centro-periferia che rispondono a macchie luminose, i neuroni della corteccia V1 hanno campi recettivi allungati che rispondono selettivamente a bordi e barre orientate in direzioni specifiche.

Questa scoperta rivoluzionaria fu fatta da David Hubel e Torsten Wiesel negli anni Sessanta attraverso esperimenti di registrazione elettrofisiologica da singoli neuroni nella corteccia visiva di gatti e scimmie. Hubel e Wiesel identificarono diversi tipi di neuroni corticali con proprietà di risposta progressivamente più complesse. Le cellule semplici hanno campi recettivi che possono essere modellati come regioni allungate con subregioni ON e OFF disposte in configurazioni parallele, rispondendo ottimalmente a bordi orientati o barre luminose in posizioni specifiche del campo visivo. Le cellule complesse rispondono a bordi

orientati ma mostrano invarianza rispetto alla posizione esatta del bordo all'interno del campo recettivo, purché l'orientazione sia corretta. Le cellule end-stopped, o ipercomplesse, rispondono a segmenti di bordo di lunghezza limitata e sono sensibili alla terminazione dei contorni.

L'organizzazione della corteccia V1 presenta una struttura geometrica notevole. I neuroni con preferenze di orientazione simili sono raggruppati in colonne verticali che attraversano tutti gli strati corticali, formando una mappa continua dell'orientazione preferita sulla superficie della corteccia. Inoltre, esistono strutture periodiche chiamate blobs, o macchie, che contengono neuroni particolarmente sensibili ai colori e meno sensibili all'orientazione. L'insieme di queste strutture forma un modulo corticale ripetuto che tesse la rappresentazione dello spazio visivo. Ogni punto della retina è rappresentato in V1 in modo retinotopico, cioè punti vicini nello spazio visivo sono rappresentati da neuroni vicini sulla superficie corticale, sebbene questa mappatura sia fortemente non uniforme, con una sovrarappresentazione massiccia della regione foveale centrale dove l'acuità visiva è massima.

Dalla corteccia V1, l'informazione visiva si propaga attraverso una gerarchia complessa di aree corticali extrastriate, ciascuna specializzata per elaborare aspetti diversi della scena visiva. La classica dicotomia proposta da Leslie Ungerleider e Mortimer Mishkin negli anni Ottanta distingue due principali vie di elaborazione: la via ventrale, o "what pathway", che si estende dalla corteccia V1 attraverso le aree V2 e V4 fino alla corteccia inferiotemporale e si specializza nel riconoscimento degli oggetti e nell'elaborazione delle forme; e la via dorsale, o "where pathway", che si estende da V1 attraverso l'area V5 (anche chiamata MT) fino alla corteccia parietale posteriore e si specializza nell'elaborazione del movimento e delle relazioni spaziali.

Lungo la via ventrale, i campi recettivi dei neuroni diventano progressivamente più grandi e le loro preferenze di stimolo diventano progressivamente più complesse e astratte. Nell'area V2, immediatamente successiva a V1, i neuroni rispondono a combinazioni più elaborate di orientazioni e iniziano a mostrare sensibilità a caratteristiche come la curvatura locale e le relazioni tra bordi adiacenti. Nell'area V4, i neuroni mostrano selettività per forme più complesse, per combinazioni specifiche di colori e orientazioni, e per configurazioni spaziali di elementi visivi. Alcuni neuroni in V4 rispondono preferenzialmente a forme geometriche come ellissi, iperboli o angoli con orientazioni specifiche.

La corteccia inferiotemporale, al culmine della via ventrale, contiene neuroni con proprietà di risposta straordinarie che furono inizialmente documentate da Charles Gross e colleghi negli anni Settanta e successivamente studiate estensivamente da molti laboratori. Alcuni neuroni in questa regione rispondono selettivamente a categorie di oggetti complessi come volti, mani, corpi o oggetti specifici. I famosi "neuroni della nonna", sebbene non rispondano letteralmente solo alla nonna di una persona, mostrano una selettività notevole per stimoli specifici, rispondendo robustamente a un particolare volto o oggetto attraverso variazioni sostanziali di posizione, scala, angolazione di visione e illuminazione. Questa invarianza della risposta neurale rispetto a trasformazioni geometriche è esattamente il tipo di comportamento che deve essere spiegato da qualsiasi teoria della percezione visiva.

Modelli Computazionali della Percezione Visiva

La neurobiologia descrittiva della via visiva, per quanto affascinante, lascia aperte questioni fondamentali riguardo ai principi computazionali che governano la trasformazione dell'informazione attraverso la gerarchia corticale e che danno origine alle proprietà di invarianza osservate sperimentalmente. Numerosi modelli computazionali sono stati proposti per spiegare come il cervello costruisca rappresentazioni invarianti degli oggetti, ciascuno enfatizzando meccanismi diversi e facendo predizioni diverse riguardo alla struttura delle rappresentazioni neurali.

Il modello HMAX (Hierarchical Model and X), sviluppato da Tomaso Poggio e colleghi al MIT, rappresenta uno degli approcci più influenti. Ispirato direttamente all'architettura gerarchica della corteccia visiva dei primati, HMAX implementa un'alternanza di operazioni di template matching e pooling attraverso molteplici stadi di elaborazione. Agli stadi bassi della gerarchia, il modello contiene unità che rispondono a gabor patches orientati, simulando le cellule semplici di V1. Queste risposte vengono poi combinate attraverso un'operazione di max-pooling su posizioni spaziali vicine, creando unità invarianti rispetto a piccole traslazioni, analoghe alle cellule complesse di V1. Questa alternanza di template matching e pooling viene ripetuta attraverso molteplici livelli, con templates progressivamente più complessi e pooling su regioni progressivamente più ampie, fino a produrre unità negli stadi alti che rispondono a caratteristiche complesse degli oggetti e mostrano invarianza sostanziale rispetto a posizione e scala.

Il modello HMAX ha mostrato prestazioni notevoli in compiti di riconoscimento di oggetti e categorizzazione rapida di scene naturali, e le sue rappresentazioni interne mostrano somiglianze significative con le rappresentazioni neurali misurate nella corteccia inferiotemporale. Tuttavia, il modello si basa fortemente su meccanismi bottom-up di elaborazione feedforward e non cattura pienamente il ruolo delle connessioni feedback che sono anatomicamente prevalenti nella corteccia visiva reale. Inoltre, sebbene il modello produca una certa invarianza, la natura esatta di questa invarianza e la sua relazione con principi matematici come la topologia rimangono poco esplorate.

Un approccio alternativo è rappresentato dalle teorie basate sulla codifica sparsa e sui dizionari adattivi. Questi modelli, influenzati da lavori pionieristici di Bruno Olshausen e David Field, propongono che la corteccia visiva apprenda un dizionario di caratteristiche elementari che permettono di rappresentare immagini naturali in modo efficiente usando solo un piccolo numero di elementi del dizionario attivi simultaneamente. La codifica sparsa ha attratto considerevole attenzione perché emerge naturalmente quando si ottimizzano rappresentazioni neurali per massimizzare l'informazione minimizzando il costo metabolico, e perché modelli di codifica sparsa producono filtri recettivi che somigliano qualitativamente ai campi recettivi misurati nella corteccia V1.

L'invarianza in questo framework emerge attraverso meccanismi di pooling e attraverso l'apprendimento di trasformazioni nel dizionario che catturano variazioni sistematiche negli stimoli. Modelli più sofisticati incorporano architetture gerarchiche di codifica sparsa dove i livelli superiori apprendono dizionari di combinazioni di elementi dei livelli inferiori, permettendo la rappresentazione di caratteristiche progressivamente più astratte e invarianti. Tuttavia, anche questi modelli generalmente non forniscono una teoria esplicita di quali proprietà invarianti vengono preservate e di come esse si relazionino a principi geometrici o topologici fondamentali.

Negli ultimi quindici anni, le reti neurali profonde e in particolare le reti neurali convoluzionali (CNN) hanno rivoluzionato il campo della visione computazionale, raggiungendo prestazioni che superano in molti compiti quelle di sistemi progettati manualmente e che in alcuni domini si avvicinano o superano le prestazioni umane. Le CNN sono architetture di deep learning che alternano strati convoluzionali, che applicano banchi di filtri appresi a patches locali dell'immagine, con strati di pooling, che riducono la risoluzione spaziale aggregando risposte su regioni vicine. Questa architettura è esplicitamente ispirata all'organizzazione della corteccia visiva e condivide molte caratteristiche strutturali con modelli come HMAX.

Studi recenti hanno rivelato somiglianze sorprendenti tra le rappresentazioni apprese dalle CNN e le rappresentazioni neurali misurate nella corteccia visiva dei primati. Utilizzando tecniche di decodifica e codifica neurale, ricercatori come Daniel Yamins e James DiCarlo hanno dimostrato che le rappresentazioni negli strati intermedi delle CNN addestrate per compiti di riconoscimento di oggetti predicono accuratamente le risposte di neuroni nella corteccia inferiotemporale a stimoli visivi complessi. Questa corrispondenza è così forte che le CNN sono ora utilizzate regolarmente come modelli computazionali della corteccia visiva, suggerendo che l'apprendimento supervisionato su grandi dataset di immagini naturali cattura almeno alcuni dei principi computazionali implementati dal cervello.

Tuttavia, sebbene le CNN mostrino robustezza impressionante rispetto a certe trasformazioni come traslazioni e, in misura limitata, cambiamenti di scala, la loro invarianza rispetto a rotazioni e deformazioni più complesse rimane imperfetta e qualitativamente diversa dall'invarianza umana. Inoltre, le CNN sono notoriamente vulnerabili ad attacchi avversariali, cioè perturbazioni impercettibili delle immagini di input che causano errori di classificazione drammatici, un fenomeno che non ha analogo chiaro nella percezione umana. Queste limitazioni suggeriscono che le CNN, nonostante i loro successi, potrebbero non catturare completamente i principi che governano la percezione visiva biologica.

È in questo contesto che l'applicazione della topologia algebrica e dell'omologia persistente offre prospettive nuove e potenzialmente trasformative. Piuttosto che focalizzarsi esclusivamente su meccanismi neurali specifici o architetture computazionali particolari, l'approccio topologico cerca di identificare le proprietà strutturali invarianti degli oggetti che il sistema visivo potrebbe estrarre e rappresentare. L'ipotesi fondamentale è che il cervello, attraverso l'evoluzione e l'apprendimento, abbia sviluppato meccanismi che sono sensibili non solo a caratteristiche geometriche metriche delle immagini ma anche, e forse soprattutto, a caratteristiche topologiche che rimangono stabili attraverso ampie classi di trasformazioni.

L'Omologia Persistente Applicata alle Immagini Naturali

Per applicare concretamente l'omologia persistente all'analisi di immagini e alla modellizzazione della percezione visiva, è necessario prima stabilire come costruire complessi simpliciali o filtrazione adatte a partire da dati di immagine. Un'immagine digitale può essere pensata come una funzione $f: \Omega \rightarrow \mathbb{R}$ dove Ω è un dominio discreto (tipicamente una griglia bidimensionale di pixel) e $f(x)$ rappresenta l'intensità luminosa o il valore cromatico al pixel x . Per estrarre informazioni topologiche da questa funzione, un approccio standard consiste nel considerare gli insiemi di livello, cioè gli insiemi di pixel dove l'intensità supera una certa soglia.

Formalmente, dato un valore di soglia t , l'insieme di livello superiore è definito come $L_t = \{x \in \Omega : f(x) \geq t\}$. Al variare della soglia t da valori alti a valori bassi, gli insiemi di livello crescono progressivamente, iniziando da piccole regioni isolate corrispondenti ai pixel più luminosi e eventualmente espandendosi fino a coprire l'intera immagine. Questa sequenza di insiemi annidati $L_{\{t_1\}} \subseteq L_{\{t_2\}} \subseteq \dots \subseteq L_{\{t_n\}}$ per $t_1 > t_2 > \dots > t_n$ costituisce una filtrazione sulla quale possiamo calcolare l'omologia persistente.

Per rendere questa procedura computazionalmente trattabile, è comune utilizzare una struttura di complesso cubico piuttosto che simpliciale. Un complesso cubico bidimensionale associato a un'immagine discreta è costruito prendendo come 0-celle i pixel, come 1-celle i lati tra pixel adiacenti, e come 2-celle i quadrati formati da quattro pixel adiacenti. Quando un pixel entra nell'insieme di livello a soglia t , anche tutte le celle di dimensione superiore incidenti su quel pixel entrano nel complesso filtrato. L'omologia del complesso cubico può essere calcolata usando algoritmi analoghi a quelli per complessi simpliciali, con il vantaggio che la struttura regolare della griglia di pixel rende le computazioni particolarmente efficienti.

Applicando l'omologia persistente a immagini naturali, si possono estrarre diagrammi di persistenza che caratterizzano le strutture topologiche presenti a diverse scale di intensità. Le componenti connesse nell'omologia di dimensione zero (H_0) corrispondono a regioni luminose separate nell'immagine. Quando la soglia diminuisce, nuove componenti appaiono in corrispondenza di nuovi massimi locali di intensità, e componenti esistenti si fondono quando le regioni luminose che rappresentano si connettono attraverso pixel di intensità intermedia. I cicli nell'omologia di dimensione uno (H_1) corrispondono a regioni circolari o anulari dove l'intensità forma un "buco" circondato da pixel più luminosi.

Studi empirici hanno dimostrato che i diagrammi di persistenza calcolati da immagini naturali contengono informazioni discriminative utili per compiti di riconoscimento e classificazione. Ad esempio, ricercatori hanno mostrato che le caratteristiche topologiche estratte da immagini di tessuti biologici permettono di distinguere tra tessuti sani e patologici con accuratezza comparabile o superiore a metodi basati su caratteristiche geometriche tradizionali. Similmente, l'omologia persistente è stata applicata con successo all'analisi di immagini mediche, immagini satellitari, immagini di materiali e molte altre tipologie di dati visivi.

Tuttavia, l'applicazione dell'omologia persistente direttamente ai valori di intensità dei pixel presenta alcune limitazioni quando l'obiettivo è modellare la percezione visiva umana. In particolare, il sistema visivo non opera direttamente sulle intensità luminose grezze ma estrae prima caratteristiche come bordi, orientazioni e texture attraverso l'elaborazione nella retina e in V1. Per questo motivo, approcci più sofisticati applicano l'omologia persistente non all'immagine originale ma a mappe di caratteristiche derivate da modelli della visione primaria.

Un approccio particolarmente promettente consiste nel calcolare l'omologia persistente sulle mappe di orientazione estratte da un'immagine. Utilizzando filtri di Gabor o altri rilevatori di bordi orientati, si può costruire per ogni posizione spaziale un istogramma di orientazioni che descrive la distribuzione locale delle orientazioni dei bordi. Queste distribuzioni di orientazioni possono essere pensate come punti in uno spazio di caratteristiche ad alta dimensione, e su questo spazio si può costruire un complesso di Vietoris-Rips o un altro tipo

di complesso simpliciale. L'omologia persistente di questo complesso cattura le relazioni topologiche tra le distribuzioni di orientazioni in diverse parti dell'immagine, fornendo una descrizione invariante rispetto a molte trasformazioni fotometriche.

Invarianza Percettiva e Evidenze Sperimentali

Teoria Matematica dell'Invarianza nelle Rappresentazioni Topologiche

La questione centrale che connette la topologia algebrica alla percezione visiva è il problema dell'invarianza: quali proprietà degli stimoli visivi rimangono stabili quando un oggetto viene osservato da angolazioni diverse, a distanze diverse, sotto illuminazioni diverse, o quando subisce deformazioni non rigide? La topologia fornisce un framework matematico naturale per formalizzare questa nozione di invarianza, poiché le proprietà topologiche sono per definizione quelle che rimangono inalterate sotto trasformazioni continue.

Per rendere questa connessione più precisa, consideriamo una classe di trasformazioni T che rappresenta le variazioni che tipicamente occorrono durante l'osservazione di oggetti tridimensionali. Questa classe potrebbe includere rotazioni tridimensionali, traslazioni, cambiamenti di scala (entro certi limiti), proiezioni prospettiche da diversi punti di vista, e deformazioni elastiche o articolate. Una rappresentazione matematica $\phi(I)$ di un'immagine I è detta invariante rispetto a T se per ogni trasformazione $t \in T$, vale la relazione $\phi(t(I)) = \phi(I)$, o più debolmente, se $\phi(t(I))$ è vicino a $\phi(I)$ secondo una qualche metrica appropriata.

Le proprietà topologiche studiate dall'omologia hanno esattamente questo carattere di invarianza rispetto a trasformazioni continue. Se due oggetti possono essere continuamente deformati l'uno nell'altro senza tagliare o incollare parti, allora hanno gli stessi gruppi di omologia e gli stessi numeri di Betti. Questo suggerisce che rappresentazioni basate sull'omologia potrebbero catturare aspetti della forma degli oggetti che sono intrinsecamente robusti rispetto alle trasformazioni che il sistema visivo deve gestire.

Tuttavia, è importante riconoscere che non tutte le trasformazioni rilevanti per la percezione visiva sono omeomorfismi nel senso stretto della topologia. Una proiezione prospettica, ad esempio, mappa uno spazio tridimensionale in uno spazio bidimensionale, riducendo la dimensionalità e introducendo necessariamente una perdita di informazione. Rotazioni tridimensionali di oggetti non convessi possono causare l'apparizione o la scomparsa di parti visibili a causa dell'occlusione, un fenomeno che non corrisponde a una deformazione continua della superficie visibile. Deformazioni articolate, come il movimento di un braccio umano, possono alterare la topologia se interpretate rigidamente, ad esempio facendo sì che un anello formato dalle braccia e dal torso si apra quando le braccia vengono abbassate.

Questi esempi suggeriscono che l'applicazione dell'omologia classica alla percezione visiva deve essere temperata da una comprensione delle limitazioni e da estensioni appropriate della teoria. L'omologia persistente offre una soluzione parziale a questo problema attraverso la sua natura multi-scala. Mentre le trasformazioni geometriche possono alterare le caratteristiche topologiche a certe scale, le caratteristiche che persistono attraverso ampie fasce di scale tendono a essere più robuste e più significative per il riconoscimento. In

questo senso, l'omologia persistente fornisce non una sola rappresentazione topologica invariante ma una famiglia di rappresentazioni a diverse scale, dalle quali il sistema di riconoscimento può estrarre le caratteristiche più stabili e discriminative.

Un risultato matematico che formalizza questa intuizione è il teorema di stabilità dell'omologia persistente menzionato in precedenza. Ricordiamo che questo teorema afferma che se due funzioni $f, g: X \rightarrow \mathbb{R}$ sono vicine nel senso della norma sup, cioè $\sup_{x \in X} |f(x) - g(x)| < \epsilon$, allora i loro diagrammi di persistenza sono vicini nel senso della distanza di bottleneck, cioè $d_B(\text{Dgm}(f), \text{Dgm}(g)) \leq \epsilon$. Questo risultato si estende a diagrammi di persistenza calcolati da complessi simpliciali costruiti su nuvole di punti attraverso una catena di stabilità che coinvolge la distanza di Hausdorff tra i dataset.

Nel contesto della percezione visiva, il teorema di stabilità ha un'interpretazione diretta: se due immagini differiscono solo per piccole perturbazioni nei valori di intensità (che potrebbero essere causate da rumore del sensore, piccole variazioni di illuminazione, o artefatti di compressione), i loro diagrammi di persistenza saranno anch'essi vicini. Questo garantisce che le caratteristiche topologiche estratte sono robuste rispetto al rumore e alle piccole variazioni non essenziali negli stimoli visivi. Tuttavia, va notato che il teorema si applica a perturbazioni che preservano la struttura generale dell'immagine e non copre trasformazioni geometriche drastiche come rotazioni di novanta gradi o ribaltamenti.

Per estendere l'invarianza topologica a classi più ampie di trasformazioni geometriche, sono necessarie tecniche aggiuntive. Un approccio consiste nell'incorporare simmetrie note esplicitamente nella costruzione del complesso simpliciale o della funzione di filtrazione. Ad esempio, se sappiamo a priori che il sistema di riconoscimento deve essere invariante rispetto a rotazioni planari, possiamo costruire descrittori di forma che sono definiti in termini di quantità rotazionalmente invarianti, come la distribuzione delle distanze tra punti o la funzione di curvatura lungo i contorni.

Un altro approccio, più recente e potenzialmente più potente, consiste nell'utilizzare l'omologia persistente equivariante, una generalizzazione dell'omologia persistente che tiene conto esplicitamente dell'azione di un gruppo di simmetrie sullo spazio dei dati. In questa formulazione, non solo si calcola l'omologia persistente ordinaria ma si traccia anche come le caratteristiche topologiche si trasformano sotto l'azione del gruppo di simmetrie. Le caratteristiche che sono invarianti rispetto al gruppo, o che si trasformano secondo rappresentazioni irriducibili note, possono essere identificate e utilizzate preferenzialmente per il riconoscimento.

Evidenze Sperimentali dall'Applicazione dell'Omologia Persistente alla Visione

Negli ultimi anni, un crescente corpo di ricerca empirica ha iniziato a esplorare se e come l'omologia persistente possa fornire insight sulla percezione visiva umana e migliorare sistemi di visione artificiale. Questi studi spaziano da analisi puramente computazionali che dimostrano l'utilità delle caratteristiche topologiche per compiti di riconoscimento, a esperimenti psicofisici che investigano se gli esseri umani sono sensibili a proprietà topologiche degli stimoli visivi, fino a studi neurofisiologici che cercano tracce di rappresentazioni topologiche nell'attività neurale del cervello.

Uno dei primi lavori significativi in questa direzione è stato condotto da Gunnar Carlsson e colleghi, che hanno applicato l'omologia persistente al problema del riconoscimento di immagini di oggetti naturali. In questo studio, le immagini sono state rappresentate non direttamente come matrici di pixel ma come punti in uno spazio di caratteristiche costruito usando piccole patches dell'immagine. Specificamente, ogni patch tre per tre di pixel è stata estratta dall'immagine e vettorizzata in un vettore di dimensione nove (o ventisette per immagini a colori). L'insieme di tutte le patches estratte da un'immagine forma una nuvola di punti in questo spazio di caratteristiche ad alta dimensione.

Su questa nuvola di punti, i ricercatori hanno costruito complessi di Vietoris-Rips e calcolato l'omologia persistente, ottenendo diagrammi di persistenza che caratterizzano la struttura topologica dello spazio delle patches. L'intuizione è che immagini di oggetti della stessa categoria, ad esempio volti umani, tendono a campionare regioni simili dello spazio delle patches, formando strutture geometriche e topologiche riconoscibili. Utilizzando tecniche di machine learning applicate ai diagrammi di persistenza, il gruppo è riuscito a classificare immagini in diverse categorie di oggetti con accuratezza significativamente superiore al caso e comparabile a metodi basati su caratteristiche geometriche tradizionali.

Questo studio pionieristico ha dimostrato la fattibilità di utilizzare l'omologia persistente come strumento per l'analisi di immagini, ma ha lasciato aperte domande su quanto precisamente questo approccio rifletta i meccanismi della percezione umana. Studi successivi hanno iniziato ad affrontare questa domanda più direttamente. Ad esempio, ricercatori presso l'Università di Chicago hanno condotto esperimenti psicofisici in cui partecipanti umani dovevano discriminare tra forme bidimensionali che differivano solo per proprietà topologiche, come il numero di buchi, mantenendo costanti caratteristiche geometriche metriche come area, perimetro e compattezza.

I risultati hanno mostrato che gli esseri umani sono infatti sensibili a proprietà topologiche e possono discriminare efficientemente tra forme con diverse topologie anche quando queste sono incorporate in configurazioni geometriche complesse che rendono difficile il riconoscimento basato su template semplici. Inoltre, la discriminazione topologica mostra caratteristiche di invarianza rispetto a deformazioni: i partecipanti identificavano correttamente il numero di buchi anche quando le forme venivano stirate, piegate o ruotate in modi che alteravano drasticamente il loro aspetto metrico. Queste osservazioni supportano l'idea che il sistema visivo umano estragga e utilizzi informazioni topologiche durante il riconoscimento delle forme.

Un altro filone di ricerca ha esplorato l'applicazione dell'omologia persistente all'analisi di texture e pattern ripetitivi. Le texture naturali, come quelle presenti sulla corteccia degli alberi, sulle superfici rocciose, sui tessuti biologici o sui materiali artificiali, spesso presentano strutture auto-simili o quasi-periodiche a diverse scale. L'omologia persistente fornisce strumenti naturali per caratterizzare queste strutture multi-scala. Studi hanno dimostrato che i diagrammi di persistenza calcolati da immagini di texture possono essere utilizzati per classificare texture in categorie, per segmentare regioni con texture diverse in immagini complesse, e per sintetizzare nuove texture che condividono le proprietà statistiche e topologiche di texture esistenti.

Nel contesto delle neuroscienze computazionali, ricercatori hanno iniziato a investigare se le rappresentazioni apprese da reti neurali artificiali contengano strutture topologiche simili a quelle che potrebbero essere presenti nel cervello. Un approccio consiste nel visualizzare lo spazio di attivazione di neuroni in uno strato intermedio di una rete neurale profonda addestrata per riconoscimento di oggetti. Quando alla rete vengono presentate immagini di oggetti appartenenti a diverse categorie, i pattern di attivazione neurale formano nuvole di punti in uno spazio ad alta dimensione. Applicando l'omologia persistente a queste nuvole, si può studiare come le diverse categorie di oggetti sono organizzate topologicamente nello spazio delle rappresentazioni.

Studi di questo tipo hanno rivelato che le categorie di oggetti semanticamente correlate tendono a formare cluster che sono topologicamente connessi, mentre categorie non correlate sono separate da regioni di bassa densità nello spazio delle attivazioni. Inoltre, la struttura topologica dello spazio delle rappresentazioni cambia attraverso gli strati della rete: strati bassi mostrano una struttura più frammentata con molte piccole componenti connesse corrispondenti a caratteristiche visive locali, mentre strati alti mostrano una struttura più coerente con poche grandi componenti connesse corrispondenti a categorie semantiche globali. Questa progressione gerarchica nella struttura topologica riflette qualitativamente l'idea che l'elaborazione visiva proceda da rappresentazioni locali e frammentate a rappresentazioni globali e integrate.

Un'applicazione particolarmente affascinante dell'omologia persistente nelle neuroscienze è lo studio delle mappe di orientazione nella corteccia visiva primaria. Come menzionato in precedenza, la corteccia V1 contiene una mappa continua delle preferenze di orientazione dove neuroni vicini rispondono preferenzialmente a orientazioni simili. Questa mappa può essere misurata sperimentalmente usando tecniche di imaging ottico o risonanza magnetica funzionale ad alta risoluzione. La mappa risultante può essere pensata come una funzione che assegna a ogni posizione sulla superficie corticale un angolo di orientazione preferito.

Sorprendentemente, queste mappe di orientazione non sono semplicemente gradienti lisci ma contengono singolarità topologiche note come punti di pinwheel, dove tutte le orientazioni convergono radialmente. Questi punti di pinwheel sono caratteristiche topologicamente robuste della mappa e il loro numero e distribuzione sono soggetti a vincoli topologici derivanti dalla teoria dell'indice di campi vettoriali su superfici. Ricercatori hanno applicato l'omologia persistente per caratterizzare quantitativamente la struttura di queste mappe e per identificare le scale a cui i punti di pinwheel rappresentano caratteristiche dominanti rispetto al rumore o alle variazioni locali.

L'analisi topologica delle mappe di orientazione ha rivelato pattern universali che sono conservati attraverso individui diversi e specie diverse, suggerendo che principi organizzativi topologici potrebbero giocare un ruolo fondamentale nello sviluppo e nella funzione della corteccia visiva. Modelli computazionali dello sviluppo corticale che incorporano vincoli topologici sono stati in grado di riprodurre quantitativamente le statistiche dei punti di pinwheel osservati sperimentalmente, fornendo supporto all'idea che la topologia sia un principio organizzativo importante a livello neurale.

Applicazioni dell'Omologia Persistente al Riconoscimento di Forme Invarianti

Oltre agli studi di base che esplorano i fondamenti teorici e le evidenze empiriche della rilevanza della topologia per la percezione visiva, un'area di ricerca applicata molto attiva riguarda lo sviluppo di algoritmi pratici di riconoscimento di forme che sfruttano l'omologia persistente per ottenere invarianza rispetto a trasformazioni. Questi algoritmi hanno trovato applicazioni in domini diversi come il riconoscimento di caratteri scritti a mano, l'identificazione di specie biologiche da immagini, la classificazione di immagini mediche, e molti altri.

Un esempio rappresentativo è il riconoscimento di caratteri manoscritti, un compito classico della visione computazionale che rimane sorprendentemente difficile a causa dell'enorme variabilità negli stili di scrittura individuali. Ogni persona scrive le lettere in modi leggermente diversi, con variazioni in scala, inclinazione, spessore del tratto, velocità di tracciamento e stile ornamentale. Tuttavia, nonostante questa variabilità metrica, certe proprietà topologiche dei caratteri tendono a rimanere stabili. La lettera "O" ha generalmente un buco, la lettera "B" ha due buchi, la lettera "A" ha un buco (se la barra trasversale è chiusa) o nessuno (se è aperta), e così via.

Ricercatori hanno sviluppato sistemi di riconoscimento che calcolano l'omologia persistente da immagini di caratteri manoscritti e utilizzano i diagrammi di persistenza risultanti come caratteristiche per la classificazione. L'approccio tipico prevede prima una fase di pre-elaborazione dove l'immagine del carattere viene binarizzata e scheletrizzata, riducendola a una curva unidimensionale che cattura la struttura essenziale del tratto. Su questa rappresentazione scheletrizzata, si calcola l'omologia persistente usando una filtrazione basata sulla distanza da un punto di riferimento o sulla direzione del gradiente dell'immagine.

I diagrammi di persistenza risultanti vengono poi vettorizzati usando uno dei metodi descritti in precedenza (paesaggi di persistenza, immagini di persistenza, o statistiche sommarie) e utilizzati come input per algoritmi di machine learning standard come macchine a vettori di supporto o reti neurali. Studi hanno dimostrato che l'inclusione di caratteristiche topologiche migliora significativamente l'accuratezza del riconoscimento, specialmente per caratteri scritti in stili non standard o degradati dal rumore.

Un altro dominio di applicazione è l'analisi di forme biologiche, dove la capacità di riconoscere organismi o strutture anatomiche indipendentemente dalla loro posa, orientazione o deformazione naturale è essenziale. Ad esempio, il riconoscimento di foglie di piante per l'identificazione automatica delle specie è un problema di grande interesse per l'ecologia e la botanica. Le foglie variano enormemente in dimensione, forma, margine e venatura, ma all'interno di una specie tendono a condividere proprietà strutturali che sono topologicamente caratterizzabili.

Sistemi di identificazione basati sull'omologia persistente hanno analizzato le reti di venature delle foglie come grafi o complessi simpliciali, calcolando caratteristiche topologiche come il numero di cicli chiusi, la distribuzione delle lunghezze dei cicli persistenti, e la struttura gerarchica della ramificazione. Queste caratteristiche topologiche, combinate con caratteristiche geometriche tradizionali, hanno permesso tassi di identificazione molto alti anche in presenza di foglie parzialmente danneggiate o di qualità di immagine subottimale.

Nel campo medico, l'omologia persistente è stata applicata all'analisi di immagini istologiche di tessuti per la diagnosi di patologie. I tessuti cancerosi, ad esempio, spesso presentano alterazioni nella loro architettura microscopica che possono essere caratterizzate topologicamente. La distribuzione spaziale dei nuclei cellulari, la struttura dei vasi sanguigni, la formazione di strutture ghiandolari o tubulari, e altri aspetti morfologici possono essere quantificati attraverso l'omologia persistente. Studi hanno mostrato che caratteristiche topologiche estratte da biopsie tissutali sono predittive della progressione del cancro e possono complementare criteri diagnostici tradizionali basati su ispezione visuale da parte di patologi esperti.

Un'applicazione particolarmente innovativa riguarda il riconoscimento di azioni umane da sequenze video. Le azioni, come camminare, correre, saltare o gesticolare, coinvolgono movimenti coordinati di parti del corpo che alterano continuamente la configurazione spaziale del corpo nel tempo. Tuttavia, certe proprietà topologiche delle configurazioni corporee possono rimanere invarianti o evolvere in modi caratteristici durante azioni specifiche. Ad esempio, durante la camminata, le gambe formano periodicamente cicli chiusi che si aprono e chiudono con la cadenza del passo.

Ricercatori hanno rappresentato sequenze video di azioni come traiettorie in spazi di caratteristiche ad alta dimensione derivati da scheletri articolati estratti automaticamente dai video. Su queste traiettorie, hanno calcolato l'omologia persistente utilizzando filtrazioni basate sul tempo o sulla distanza nello spazio delle pose. I diagrammi di persistenza risultanti catturano pattern temporali e spaziali delle azioni che sono robusti rispetto a variazioni individuali nello stile di esecuzione, nella velocità, e nell'orientazione relativa rispetto alla camera. Sistemi di riconoscimento basati su queste rappresentazioni topologiche hanno mostrato prestazioni competitive con approcci state-of-the-art basati su deep learning, con il vantaggio aggiunto di maggiore interpretabilità e minore dipendenza da grandi quantità di dati di addestramento.

Teorie Avanzate, Sfide Aperte e Prospettive Future

Geometria Differenziale e Topologia nella Rappresentazione delle Forme

Mentre l'omologia persistente fornisce un framework potente per estrarre caratteristiche topologiche da dati visivi, un'area di ricerca complementare esplora come la geometria differenziale possa essere integrata con la topologia algebrica per costruire descrittori di forma ancora più ricchi e discriminativi. La geometria differenziale studia le proprietà di curve, superfici e varietà che dipendono da concetti di derivazione e curvatura, fornendo strumenti per caratterizzare la forma locale degli oggetti in modi che vanno oltre le distinzioni puramente topologiche.

Un concetto centrale nella geometria differenziale delle superfici è la curvatura, che misura quanto una superficie si discosta localmente dall'essere piatta. Per superfici bidimensionali immerse nello spazio tridimensionale, esistono due quantità fondamentali: la curvatura gaussiana K , definita come il prodotto delle curvature principali, e la curvatura media H , definita come la media aritmetica delle curvature principali. La curvatura gaussiana ha la proprietà notevole di essere un invariante intrinseco della superficie, cioè dipende solo dalla

metrica interna della superficie e non dal modo in cui essa è immersa nello spazio ambiente. Questo significa che se una superficie viene piegata senza stirarsi, la sua curvatura gaussiana rimane invariata.

Il teorema di Gauss-Bonnet fornisce una connessione profonda tra la geometria differenziale e la topologia, stabilendo che l'integrale della curvatura gaussiana su una superficie compatta chiusa è uguale a 2π per la caratteristica di Eulero della superficie: $\iint_S K \, dA = 2\pi\chi(S)$, dove $\chi(S)$ è la caratteristica di Eulero, definita come $\chi(S) = 2 - 2g$ per una superficie di genere g (cioè con g buchi). Questo teorema straordinario mostra che una quantità puramente geometrica (l'integrale della curvatura) è determinata da una quantità puramente topologica (la caratteristica di Eulero), costruendo un ponte tra le due discipline matematiche.

Nel contesto della percezione visiva, la distribuzione della curvatura su una superficie fornisce informazioni cruciali sulla forma tridimensionale degli oggetti. Il sistema visivo umano è notevolmente sensibile alla curvatura, come dimostrato da numerosi studi psicofisici. Osservatori possono discriminare differenze sottili nella curvatura di superfici anche quando queste sono presentate con cue minimali di profondità. Inoltre, la segmentazione percettiva delle superfici in parti significative sembra seguire discontinuità nella curvatura: i confini tra parti percettivamente distinte spesso corrispondono a creste o valli nella distribuzione di curvatura, cioè a curve lungo le quali la curvatura cambia drasticamente.

Ricercatori hanno sviluppato descrittori di forma che combinano informazioni topologiche e geometriche calcolando l'omologia persistente non sull'oggetto grezzo ma su funzioni di curvatura definite sull'oggetto. Ad esempio, data una superficie triangolata, si può calcolare la curvatura gaussiana o media a ogni vertice del mesh usando formule discrete appropriate. Questa funzione di curvatura può poi essere utilizzata come funzione di filtrazione per l'omologia persistente. Il diagramma di persistenza risultante caratterizza le scale a cui caratteristiche topologiche della funzione di curvatura emergono e scompaiono, fornendo una firma multi-scala che codifica sia la struttura topologica che quella geometrica della superficie.

Alternativamente, si possono costruire filtrazioni basate su insiemi di livello della curvatura. Per una superficie S con funzione di curvatura $\kappa: S \rightarrow \mathbb{R}$, l'insieme di livello superiore $S_t = \{p \in S : \kappa(p) \geq t\}$ consiste di regioni della superficie dove la curvatura eccede la soglia t . Al variare di t , questi insiemi formano una filtrazione sulla quale si può calcolare l'omologia persistente. Le componenti connesse nell'omologia H_0 corrispondono a regioni di alta curvatura separate (come cuspidi, creste o protuberanze), mentre i cicli nell'omologia H_1 corrispondono a regioni anulari di alta curvatura (come creste circolari o valli). La persistenza di queste caratteristiche riflette quanto robustamente esse emergono rispetto al livello di soglia di curvatura.

Un framework teorico particolarmente elegante per unificare topologia e geometria nella rappresentazione delle forme è fornito dalla teoria di Morse, sviluppata dal matematico Marston Morse negli anni Trenta del secolo scorso. Una funzione di Morse è una funzione liscia $f: M \rightarrow \mathbb{R}$ su una varietà differenziabile M che ha solo punti critici non degeneri, cioè punti dove il gradiente si annulla e la matrice hessiana è non singolare. Ogni punto critico ha

un indice che corrisponde al numero di autovalori negativi dell'hessiana: un minimo locale ha indice zero, un punto sella ha indice uno su una superficie bidimensionale, e un massimo locale ha indice due.

La teoria di Morse stabilisce una relazione profonda tra i punti critici di una funzione e la topologia della varietà sottostante. Quando una funzione di Morse viene utilizzata come funzione di filtrazione per l'omologia persistente, i punti critici corrispondono esattamente ai punti dove caratteristiche topologiche nascono o muoiono nella filtrazione. Specificamente, quando si attraversa un punto critico di indice k aumentando il parametro di filtrazione, o nasce un nuovo k -ciclo o muore un $(k-1)$ -ciclo. Questa corrispondenza permette di calcolare l'omologia persistente in modo efficiente per funzioni di Morse e fornisce un'interpretazione geometrica chiara degli intervalli di persistenza in termini di coppie critiche di punti.

Nel contesto della percezione visiva, le funzioni di Morse naturali includono la funzione altezza (che misura l'elevazione di punti sulla superficie rispetto a un piano di riferimento), la funzione di distanza da un punto o da un insieme di riferimento, e funzioni intrinseche derivate dalla geometria della superficie come la funzione di distanza geodetica da un punto seme. Ognuna di queste funzioni cattura aspetti diversi della forma e produce diagrammi di persistenza che enfatizzano caratteristiche strutturali diverse.

Teoria delle Categorie e Funtorialità dell'Omologia Persistente

Un'area di ricerca matematica più astratta ma potenzialmente molto fruttuosa per la comprensione teorica della percezione visiva è l'applicazione della teoria delle categorie all'omologia persistente e all'analisi topologica dei dati. La teoria delle categorie fornisce un linguaggio universale per descrivere strutture matematiche e le relazioni tra di esse, e ha trovato applicazioni sorprendentemente ampie in diverse aree della matematica, della fisica teorica e, più recentemente, dell'informatica teorica e della scienza dei dati.

In questo framework, l'omologia persistente può essere vista come un funtore dalla categoria degli spazi topologici filtrati alla categoria dei moduli persistenti. Un funtore è una mappa tra categorie che preserva la struttura compositiva: esso associa a ogni oggetto di una categoria un oggetto dell'altra categoria, e a ogni morfismo un morfismo, in modo che identità e composizioni siano preservate. La proprietà funtoriale dell'omologia persistente significa che relazioni tra spazi filtrati (come inclusioni o mappe continue) inducono relazioni tra i corrispondenti moduli persistenti in modo coerente e compositivo.

Questa prospettiva categoriale ha implicazioni profonde per come pensiamo all'invarianza nella percezione visiva. Se consideriamo l'insieme di tutte le possibili immagini retiniche come oggetti di una categoria, e le trasformazioni percettivamente irrilevanti (come piccole rotazioni, traslazioni, cambiamenti di scala) come morfismi in questa categoria, allora una buona rappresentazione per il riconoscimento dovrebbe comportarsi funtorialmente, mappando immagini in rappresentazioni e trasformazioni in relazioni tra rappresentazioni in modo che la struttura compositiva sia preservata. L'omologia persistente, essendo un funtore, soddisfa automaticamente questa proprietà.

Inoltre, la teoria delle categorie fornisce strumenti per costruire nuovi funtori componendo funtori esistenti. Questo suggerisce un approccio modulare alla costruzione di pipeline di elaborazione visiva: ogni stadio di elaborazione (estrazione di bordi, calcolo di

caratteristiche, pooling spaziale, ecc.) può essere visto come un funtore che trasforma una categoria di rappresentazioni in un'altra, e l'intera pipeline è la composizione di questi funtori. La funtorialità garantisce che proprietà di invarianza a livello locale (cioè a livello di singoli stadi) si propaghino attraverso la pipeline, producendo invarianza globale nel sistema completo.

Un concetto particolarmente interessante dalla teoria delle categorie applicata all'omologia persistente è quello di interleaving distance, una metrica tra moduli persistenti introdotta da Michael Lesnick e Dmitriy Morozov. Questa distanza generalizza la distanza di bottleneck tra diagrammi di persistenza e può essere definita anche per moduli persistenti più generali che non ammettono una decomposizione semplice in intervalli. La interleaving distance ha proprietà teoriche attraenti: è una pseudometrica (soddisfa disuguaglianza triangolare e simmetria), è stabile rispetto a perturbazioni, e cattura in modo naturale la nozione di quanto due rappresentazioni persistenti siano simili.

Nel contesto della percezione visiva, la interleaving distance potrebbe fornire una base teorica per definire quando due oggetti dovrebbero essere considerati percettivamente simili dal punto di vista topologico. Due oggetti i cui moduli persistenti hanno piccola interleaving distance condividono strutture topologiche multi-scala simili e potrebbero ragionevolmente essere confusi o categorizzati insieme dal sistema visivo. Al contrario, oggetti con grande interleaving distance tra i loro moduli persistenti hanno strutture topologiche fondamentalmente diverse e dovrebbero essere facilmente distinguibili.

Integrazione con il Deep Learning e Reti Neurali Topologiche

Una delle direzioni di ricerca più promettenti e attivamente perseguite è l'integrazione di concetti e metodi dalla topologia algebrica con architetture di deep learning. Mentre le reti neurali profonde hanno dimostrato capacità straordinarie nel riconoscimento visivo, esse tipicamente operano come "scatole nere" le cui rappresentazioni interne sono difficili da interpretare e i cui principi operativi rimangono largamente opachi. L'integrazione di strutture topologiche esplicite nelle architetture neurali potrebbe sia migliorare le prestazioni che aumentare l'interpretabilità dei modelli.

Un approccio consiste nell'utilizzare l'omologia persistente come uno strato di elaborazione all'interno di una rete neurale profonda. Piuttosto che applicare l'omologia persistente solo ai dati di input o agli output finali, si può calcolare l'omologia persistente sulle mappe di caratteristiche prodotte da strati intermedi della rete. Questo richiede lo sviluppo di implementazioni differenziabili dell'omologia persistente che permettano il calcolo dei gradienti attraverso le operazioni topologiche, in modo che l'intera rete possa essere addestrata end-to-end usando backpropagation.

Recenti lavori hanno fatto progressi significativi in questa direzione. Ricercatori hanno sviluppato layer topologici differenziabili che approssimano il calcolo di caratteristiche persistenti in modi che sono compatibili con l'ottimizzazione basata su gradienti. Ad esempio, il persistence codebook layer calcola una rappresentazione vettoriale approssimata del diagramma di persistenza usando funzioni soft-min e soft-max che sono lisce e differenziabili, permettendo ai gradienti di fluire attraverso lo strato durante l'addestramento. Altri approcci utilizzano rilassamenti continui della costruzione del

complesso simpliciale o approssimazioni neurali delle mappe di vettorizzazione dai diagrammi di persistenza ai vettori di caratteristiche.

L'integrazione di layer topologici in reti neurali profonde ha mostrato benefici in diversi contesti applicativi. In particolare, per compiti dove l'invarianza rispetto a deformazioni complesse è critica, come il riconoscimento di forme deformabili o la segmentazione di strutture anatomiche in immagini mediche, l'inclusione di caratteristiche topologiche esplicite può migliorare significativamente la robustezza e l'accuratezza del modello. Inoltre, i layer topologici possono servire come meccanismi di regolarizzazione, incoraggiando la rete ad apprendere rappresentazioni che preservano strutture topologiche rilevanti e ignorano variazioni irrilevanti.

Un'altra direzione di ricerca esplora l'uso di grafi e complessi simpliciali come domini nativi per operazioni di convoluzione neurale, generalizzando le reti neurali convoluzionali tradizionali che operano su griglie regolari. Le graph neural networks (GNN) e le loro estensioni a complessi simpliciali, note come simplicial neural networks, permettono di processare dati che hanno strutture topologiche intrinseche complesse, come reti sociali, molecole, mesh tridimensionali o grafi di scena visiva. In queste architetture, le operazioni di convoluzione sono definite in termini di messaggi scambiati tra nodi adiacenti del grafo o tra simplessi adiacenti del complesso, e possono incorporare informazioni topologiche come il numero di cicli locali o la struttura di connessione ad alta dimensione.

Per applicazioni alla percezione visiva, si possono costruire grafi o complessi simpliciali a partire da immagini utilizzando tecniche diverse. Un approccio consiste nel segmentare l'immagine in regioni (superpixels) e costruire un grafo dove i nodi rappresentano regioni e gli spigoli connettono regioni adiacenti. Su questo grafo di regione, una graph neural network può propagare informazioni tra regioni vicine, raffinando progressivamente le rappresentazioni in modo che catturino relazioni spaziali e semantiche complesse. Alternativamente, si possono costruire complessi simpliciali direttamente dalla geometria intrinseca dell'immagine, utilizzando metodi come la triangolazione di Delaunay o il complesso alpha sui pixel o sui keypoint estratti.

Limiti Teorici e Sfide Computazionali

Nonostante i progressi sostanziali negli ultimi anni, l'applicazione della topologia algebrica alla percezione visiva affronta diverse sfide teoriche e computazionali che richiedono ulteriori ricerche per essere superate completamente. Una limitazione fondamentale dell'omologia persistente classica è che essa cattura solo aspetti relativamente grezzi della struttura topologica, specificamente il numero e la persistenza di componenti connesse e cicli. Molte distinzioni di forma sottili che sono percettivamente rilevanti non sono catturate dai numeri di Betti o dai diagrammi di persistenza standard.

Ad esempio, consideriamo due superfici toroidali, entrambe con genere uno (un buco). Dal punto di vista dell'omologia classica, questi due tori sono indistinguibili: entrambi hanno $\beta_0 = 1$, $\beta_1 = 2$, $\beta_2 = 1$. Tuttavia, i tori potrebbero avere forme metriche molto diverse: uno potrebbe essere un toro sottile e allungato simile a un tubo piegato, mentre l'altro potrebbe essere un toro grassoccio quasi sferico con un piccolo buco. Queste differenze geometriche

sono percettivamente salienti e rilevanti per il riconoscimento, ma l'omologia da sola non le cattura.

Per affrontare questa limitazione, ricercatori hanno esplorato estensioni e generalizzazioni dell'omologia che codificano informazioni geometriche più ricche. Un approccio consiste nell'utilizzare omologia persistente multi-parametro, dove invece di una singola funzione di filtrazione parametrizzata da un numero reale, si utilizzano filtrazioni parametrizzate da vettori multidimensionali. Ad esempio, si potrebbe filtrare simultaneamente rispetto alla curvatura gaussiana e alla curvatura media, ottenendo una famiglia di complessi indicizzati da coppie (t_1, t_2) nel piano. L'omologia persistente multi-parametro è teoricamente molto più ricca dell'omologia persistente a un parametro, ma è anche computazionalmente molto più complessa e non ammette in generale una decomposizione semplice in moduli intervallo.

Un'altra direzione è l'utilizzo di invarianti topologici più sofisticati oltre all'omologia, come la coomologia, i gruppi di omotopia, o invarianti derivati dalla topologia algebrica superiore. La coomologia, che è algebricamente duale all'omologia, codifica informazioni su come cicli possono essere "colle" insieme e ha strutture aggiuntive come il cup product che non hanno analoghi nell'omologia. I gruppi di omotopia, che classificano mappe continue da sfere nella varietà, catturano aspetti della struttura topologica che vanno oltre la semplice connessione e il numero di buchi. Tuttavia, questi invarianti sono generalmente molto più difficili da calcolare algebricamente e la loro applicazione pratica all'analisi di dati rimane limitata.

Dal punto di vista computazionale, una sfida maggiore è la scalabilità degli algoritmi di omologia persistente a dataset di grandi dimensioni. Come menzionato in precedenza, il numero di semplici in un complesso di Vietoris-Rips o Čech può crescere esponenzialmente con il numero di punti e la dimensione massima dei semplici considerati. Per dataset con decine di migliaia di punti, che sono comuni in applicazioni moderne, il calcolo dell'omologia persistente in dimensioni elevate diventa rapidamente impraticabile con algoritmi standard.

Diverse strategie sono state sviluppate per mitigare questo problema. Gli algoritmi di approssimazione calcolano versioni approssimate dell'omologia persistente con garanzie teoriche sulla vicinanza all'omologia esatta, sacrificando precisione per velocità. Le tecniche di downsampling riducono il numero di punti nel dataset selezionando subset rappresentativi attraverso metodi come landmark selection o core set construction. Gli approcci basati su sparse filtrations costruiscono complessi simpliciali ridotti che contengono solo un sottoinsieme dei semplici possibili, scelti strategicamente per preservare le caratteristiche topologiche più significative minimizzando la complessità computazionale.

Un'altra sfida è la sensibilità dell'omologia persistente alla scelta di metriche e parametri. La costruzione di complessi simpliciali richiede di definire una nozione di distanza o similarità tra punti, e diverse scelte di metrica possono produrre diagrammi di persistenza molto diversi. Similmente, la scelta di quali dimensioni omologiche calcolare, quali intervalli di scale considerare, e come vettorizzare i diagrammi per il machine learning introduce gradi di libertà che devono essere fissati attraverso validazione empirica o conoscenza del dominio. Non esiste una procedura universalmente ottimale per fare queste scelte, e in pratica una quantità considerevole di sperimentazione è spesso necessaria.

Connessioni con la Teoria dell'Informazione e la Compressione

Un framework concettuale complementare per comprendere la percezione visiva è fornito dalla teoria dell'informazione, che studia la quantificazione, lo storage e la comunicazione dell'informazione. Una prospettiva influente, articolata da Horace Barlow e altri teorici della visione negli anni Sessanta e Settanta, è che il sistema visivo sia ottimizzato per estrarre e rappresentare informazione in modo efficiente, codificando stimoli naturali usando il minor numero possibile di bit pur preservando le informazioni rilevanti per il comportamento.

Questa prospettiva di codifica efficiente può essere connessa alla topologia algebrica attraverso l'idea che le caratteristiche topologiche persistenti rappresentano forme compresse di informazione strutturale. Un diagramma di persistenza con poche caratteristiche ad alta persistenza fornisce una descrizione molto compatta della struttura topologica multi-scala di un dataset, richiedendo molti meno bit per essere memorizzato rispetto ai dati grezzi originali. Allo stesso tempo, se le caratteristiche topologiche sono scelte appropriatamente, questa rappresentazione compressa può preservare le informazioni necessarie per compiti di riconoscimento e categorizzazione.

Questa connessione può essere formalizzata usando concetti dalla lossy compression e rate-distortion theory. In questo framework, definiamo una funzione di distorsione $D(x, \hat{x})$ che misura quanto una rappresentazione compressa $\hat{x}$ differisce dai dati originali x . L'obiettivo è trovare uno schema di codifica che minimizzi il rate (il numero di bit necessari per rappresentare $\hat{x}$) soggetto a un vincolo sulla distorsione massima ammissibile. Per dati visivi, potremmo definire la distorsione in termini di distanza tra diagrammi di persistenza o tra altre caratteristiche percettivamente rilevanti, piuttosto che in termini di errore quadratico medio sui pixel.

Ricerche preliminari in questa direzione hanno esplorato l'uso dell'omologia persistente per compressione di forme tridimensionali. L'idea è di rappresentare una superficie complessa attraverso il suo diagramma di persistenza insieme a informazioni geometriche minime necessarie per ricostruire approssimativamente la superficie. Poiché il diagramma di persistenza è tipicamente molto più compatto della rappresentazione completa della superficie (che potrebbe richiedere migliaia o milioni di triangoli in un mesh), questo approccio può ottenere fattori di compressione significativi. La superficie ricostruita potrebbe non essere identica all'originale a livello di dettagli fini, ma preserva le caratteristiche topologiche e geometriche principali che sono percettivamente più salienti.

Prospettive Future e Direzioni di Ricerca

Guardando al futuro, numerose direzioni di ricerca promettenti si aprono all'intersezione tra topologia algebrica, neuroscienze e visione computazionale. Sul fronte teorico, rimangono molte domande fondamentali aperte riguardo a quali esattamente siano le rappresentazioni topologiche ottimali per diversi compiti percettivi, come queste rappresentazioni possano essere apprese efficientemente da dati sensoriali, e come esse si relazionino ad altre teorie della percezione basate su principi bayesiani, di codifica predittiva, o di minimizzazione dell'energia libera.

Una direzione particolarmente eccitante è l'integrazione di topologia algebrica con modelli probabilistici e inferenza bayesiana. Mentre l'omologia persistente fornisce descrittori deterministici di strutture topologiche, molti problemi percettivi sono intrinsecamente probabilistici a causa di rumore sensoriale, ambiguità e variabilità naturale negli stimoli. Sviluppare framework bayesiani dove i parametri del modello includono caratteristiche topologiche, e dove l'inferenza permette di stimare distribuzioni posteriori su spazi di forme topologicamente vincolati, potrebbe fornire modelli più realistici e potenti della percezione umana.

Sul fronte sperimentale, sono necessari studi neurofisiologici più diretti che investigino se e come rappresentazioni topologiche siano implementate nel cervello biologico. Mentre esistono evidenze indirette da esperimenti psicofisici e da analisi computazionali di attività neurale registrata, mancano ancora dimostrazioni dirette di neuroni o popolazioni neurali che codificano esplicitamente caratteristiche topologiche come numeri di Betti o intervalli di persistenza. Tecniche avanzate di imaging neurale ad alta risoluzione, combinato con paradigmi sperimentali progettati specificamente per testare predizioni della teoria topologica, potrebbero colmare questa lacuna.

Dal punto di vista applicativo, l'integrazione sempre più stretta di metodi topologici con sistemi di visione artificiale operativi potrebbe portare a miglioramenti sostanziali in robustezza, generalizzazione e interpretabilità. Applicazioni particolarmente promettenti includono la robotica autonoma, dove robot devono riconoscere e manipolare oggetti in ambienti non strutturati sotto condizioni di visualizzazione variabili; la realtà aumentata, dove sistemi devono tracciare e riconoscere oggetti del mondo reale in tempo reale; e la diagnostica medica assistita da computer, dove l'analisi automatica di immagini deve essere non solo accurata ma anche spiegabile e affidabile.

Infine, una direzione di ricerca a lungo termine riguarda lo sviluppo di teorie unificate della percezione che integrino prospettive topologiche, geometriche, statistiche e computazionali in un framework coerente. La percezione visiva è un fenomeno multisfaccettato che coinvolge elaborazione di informazione a molteplici scale temporali e spaziali, dall'attività di singoli fotorecettori alla dinamica di grandi reti corticali, dall'estrazione di caratteristiche locali all'inferenza di scene globali, dall'elaborazione bottom-up guidata dagli stimoli all'influenza top-down delle aspettative e del contesto. Una teoria completa deve catturare tutti questi aspetti e le loro interazioni.

La topologia algebrica, con il suo focus su proprietà strutturali invarianti e la sua capacità di operare attraverso scale multiple, offre strumenti preziosi per questa impresa teorica ambiziosa. Tuttavia, essa rappresenta solo un pezzo del puzzle più grande. Il vero progresso verrà probabilmente dall'integrazione creativa di idee dalla topologia con concetti dalla geometria differenziale, dalla teoria dell'informazione, dalla statistica, dall'ottimizzazione, dalle neuroscienze sperimentali e dalla psicologia cognitiva. È attraverso questo dialogo interdisciplinare che possiamo sperare di svelare i principi computazionali che permettono al cervello di computare forme e riconoscere il mondo visivo con la straordinaria facilità e robustezza che caratterizza la percezione umana.

Conclusioni: Verso una Comprensione Topologica della Percezione

In questo saggio esteso, abbiamo esplorato come la topologia algebrica, e in particolare l'omologia persistente, fornisca un framework matematico potente per comprendere come il cervello computa forme e riconosce oggetti invarianti rispetto a trasformazioni. Abbiamo visto che le proprietà topologiche, essendo per loro natura robuste rispetto a deformazioni continue, catturano aspetti strutturali degli oggetti che rimangono stabili attraverso le variazioni geometriche che tipicamente occorrono durante l'osservazione visiva.

L'omologia persistente estende l'omologia classica fornendo descrizioni multi-scala della struttura topologica, permettendo di identificare caratteristiche che persistono attraverso ampie gamme di scale come segnali robusti distinti dal rumore transitorio. Il teorema di stabilità garantisce che queste caratteristiche sono insensibili a piccole perturbazioni, una proprietà essenziale per applicazioni a dati reali rumorosi. Attraverso tecniche di vettorizzazione come paesaggi di persistenza e immagini di persistenza, i diagrammi di persistenza possono essere integrati in pipeline di machine learning moderne, combinando i vantaggi dell'invarianza topologica con la potenza dell'apprendimento statistico.

Le evidenze empiriche, sebbene ancora in sviluppo, suggeriscono che gli esseri umani sono sensibili a proprietà topologiche degli stimoli visivi e che caratteristiche topologiche possono migliorare le prestazioni di sistemi di riconoscimento artificiale in compiti dove l'invarianza rispetto a deformazioni è critica. L'integrazione di layer topologici in architetture di deep learning rappresenta una direzione promettente per costruire sistemi di visione che combinano la capacità di apprendimento delle reti neurali con i principi di invarianza strutturale della topologia. Tuttavia, rimangono molte sfide da affrontare. Dal punto di vista teorico, dobbiamo sviluppare una comprensione più profonda di come le rappresentazioni topologiche si relazionino ad altri aspetti della percezione visiva, come la codifica del colore, la percezione del movimento, e l'integrazione di informazioni da diverse modalità sensoriali. Dal punto di vista computazionale, sono necessari algoritmi più efficienti e scalabili per calcolare caratteristiche topologiche da dataset molto grandi e ad alta dimensione. Dal punto di vista neuroscientifico, serve maggiore evidenza diretta che il cervello implementi effettivamente computazioni che possono essere interpretate in termini topologici. Nonostante queste sfide, il programma di ricerca che applica la topologia algebrica alla percezione visiva ha già prodotto insight significativi e risultati pratici in poco più di due decenni di sviluppo intensivo. La convergenza di matematica pura, neuroscienze computazionali, visione artificiale e machine learning in questo dominio esemplifica il potere dell'approccio interdisciplinare nella scienza cognitiva contemporanea. Man mano che gli strumenti matematici diventano più raffinati, le tecniche computazionali più efficienti, e le evidenze empiriche più complete, possiamo aspettarci che la topologia giochi un ruolo sempre più centrale nella nostra comprensione di come il cervello percepisce e rappresenta il mondo visivo. Il viaggio dal fotone che colpisce un fotorecettore alla rappresentazione mentale consapevole di un oggetto riconosciuto è straordinariamente complesso, coinvolgendo miliardi di neuroni e innumerevoli operazioni computazionali distribuite attraverso dozzine di aree cerebrali. Tuttavia, attraverso l'applicazione di principi matematici astratti come quelli della topologia algebrica, stiamo iniziando a intravedere i principi organizzativi che governano questa trasformazione. Forse, in definitiva, la capacità del cervello di riconoscere una tazza come tazza indipendentemente da come la guardiamo riflette non solo sofisticati meccanismi neurali ma anche verità matematiche profonde sulla natura della forma, dello spazio e dell'invarianza strutturale che trascendono il substrato biologico particolare della cognizione umana.